

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
30 octobre 2003 (30.10.2003)

PCT

(10) Numéro de publication internationale  
**WO 03/088979 A2**

(51) Classification internationale des brevets<sup>7</sup> :  
**A61K 31/7088**, 38/02, 39/21, C07D 15/15, C07K  
16/10, C12N 15/48, C12Q 1/68, 1/70

(74) Mandataires : **DEMACHY, Charles** etc.; Gros-  
set-Fournier & Demachy, 54, rue Saint-Lazare, F-75009  
Paris (FR).

(21) Numéro de la demande internationale :  
PCT/FR03/01274

(81) États désignés (*national*) : AE, AG, AL, AM, AT, AU, AZ,  
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,  
DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,  
MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE,  
SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,  
VC, VN, YU, ZA, ZM, ZW.

(22) Date de dépôt international : 22 avril 2003 (22.04.2003)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :  
02/05001 22 avril 2002 (22.04.2002) FR

(84) États désignés (*régional*) : brevet ARIPO (GH, GM, KE,  
LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet  
eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet  
européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,  
FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK,  
TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, ML, MR, NE, SN, TD, TG).

(71) Déposants (*pour tous les États désignés sauf US*) :  
**CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE** [FR/FR]; 3, rue Michel Ange, F-75794 Paris  
Cedex 16 (FR). **UNIVERSITE DE MONTPELLIER II**  
[FR/FR]; 2 place Eugène Bataillon, F-34095 MONTPEL-  
LIER Cedex 5 (FR).

**Publiée :**

— *sans rapport de recherche internationale, sera republiée  
dès réception de ce rapport*

(72) Inventeurs; et

(75) Inventeurs/Déposants (*pour US seulement*) : **KIM, Félix  
Jinhyun** [FR/FR]; 65, avenue du Pont Juvenal, Apt.69,  
F-34000 Montpellier (FR). **MANEL, Nicolas Gabriel Al-  
bert** [FR/FR]; 11, rue Robert Desnos, F-34070 Montpellier  
(FR). **SITBON, Marc Khamous Michel** [FR/FR]; 17, rue  
de Louvain, F-34000 Montpellier (FR).

*En ce qui concerne les codes à deux lettres et autres abrégia-  
tions, se référer aux "Notes explicatives relatives aux codes et  
abréviations" figurant au début de chaque numéro ordinaire de  
la Gazette du PCT.*

(54) Title: OLIGONUCLEOTIDES FROM SEQUENCES CODING FOR THE SURFACE COMPONENT OF PTLV ENVELOPE  
PROTEINS AND USES THEREOF

(54) Titre : OLIGONUCLEOTIDE ISSUS DES SEQUENCES CODANT POUR LA COMPOSANTE DE SURFACE DES PRO-  
TEINES D'ENVELOPPE DES PTLV ET LEURS UTILISATIONS

(57) Abstract: The invention relates to the use of oligonucleotides from the nucleotide sequences coding for the amino-terminal  
region of the surface component (SU) of envelope proteins of PTLV viruses in order to perform methods of detecting every PTLV  
strain or PTLV-related viruses, e.g. for the detection of novel PTLV variants or viruses comprising sequences related to PTLV SUs.  
The invention also relates to primer pairs which are used to perform said detection methods and the novel PTLV variants thus detected.

(57) Abrégé : L'invention a pour objet l'utilisation d'oligonucléotides issus des séquences nucléotiques codant pour la région ami-  
noterminal de la composante de surface (SU) des protéines d'enveloppe des virus des PTLV, pour la mise en oeuvre de procédés de  
détection de toute souche de PTLV, ou de virus apparentés aux PTLV, notamment pour la détection de nouveaux variants des PTLV,  
ou de virus comportant des séquences apparentées aux SU des PTLV. L'invention a également pour objet des couples d'amorces pour  
la mise en oeuvre de ces procédés de détection, ainsi que les nouveaux variants de PTLV ainsi détectés.



WO 03/088979 A2

## OLIGONUCLEOTIDES ISSUS DES SEQUENCES CODANT POUR LA COMPOSANTE DE SURFACE DES PROTEINES D'ENVELOPPE DES PTLV ET LEURS UTILISATIONS

---

5

La présente invention a pour objet des oligonucléotides issus des séquences nucléotidiques codant pour la région aminoterminal de la composante de surface des protéines d'enveloppe des virus des lymphomes/leucémies T chez les primates, regroupés sous la désignation PTLV, et leurs utilisations dans le cadre de la détection de toute souche de PTLV ou de souches virales apparentées.

La présente invention découle de l'identification par les Inventeurs de motifs peptidiques de la SU qui conviennent à la synthèse d'oligonucléotides pouvant être utilisés pour la détection et l'amplification de séquences pan-PTLV comprenant ces motifs. Les inventeurs ont mis au point une méthode permettant l'amplification de telles séquences, leur clonage et séquençage. La présente invention permet notamment la détection de séquences individuelles présentes dans un mélange de séquences des différents types. L'optimisation pour certains des motifs peptidiques ainsi identifiés a déjà permis la caractérisation de variants PTLV jamais encore décrits, ainsi que de détecter des séquences PTLV dont la présence dans les échantillons testés n'était pas suspectée. L'application généralisée de la présente invention permettra la détection et la caractérisation soit de nouvelles séquences apparentées aux SU de PTLV, soit de séquences déjà connues dans de nouveaux contextes pathologiques ou non.

La recherche de séquences des rétrovirus humains ou de primates est primordiale dans de nombreux contextes. De manière non-exhaustive, ces recherches intéressent le criblage de matériels biologiques (produits dérivés du sang, par exemple), le diagnostic (recherche de l'étiologie de syndromes multiples couvrant leucémies, maladies dégénératives, maladies autoimmunes, etc), les études épidémiologiques et anthropologiques des différents groupes humains, le séquençage des génomes (composition et marqueurs rétroviraux polymorphiques des génomes), le criblage de nouveaux médicaments (définition de nouvelles cibles), etc.

Dans le cas des PTLV, nous relèverons deux exemples des problèmes associés à la détection de leurs séquences. Dans le premier exemple, des individus, généralement réunis sous le terme de "séroindéterminés", présentent une réponse immune anti-HTLV dite "incomplète", dirigée contre certains antigènes seulement des PTLV, alors qu'aucune séquence correspondant à des PTLV ne peut être amplifiée à partir d'échantillons sanguins de

ces patients. Dans les cas les mieux documentés la recherche de telles séquences se fait sur des régions conservées des gènes *gag*, *pol*, *env* et *tax*. Dans le cas du gène d'enveloppe, la partie amino terminale de la SU est écartée de cette approche du fait de sa variabilité. La région amino terminale, de la composante de surface (SU) des enveloppes des rétrovirus de primates humains et non-humains de type HTLV et STLV (regroupés ici sous le terme PTLV) est notamment responsable de la reconnaissance du ou des récepteurs cellulaires pour l'enveloppe (Kim et coll, 2000). À ce jour, aucune méthode d'amplification dans cette région directement applicable aux trois types de PTLV (application dite pan-PTLV) n'a été décrite. Ainsi, généralement seule est considérée l'amplification de motifs présents dans les parties les plus conservées de la composante transmembranaire de l'enveloppe (TM). Toutefois, dans la mesure où la variabilité de la SU est un élément essentiel de la biologie adaptative des rétrovirus, la mise au point d'une approche basée sur sa détection représente un objectif particulièrement intéressant.

La présente invention a pour objet l'utilisation de couples d'oligonucléotides dégénérés en orientation 5' et 3' issus des séquences nucléotidiques codant pour la région aminoterminal de la composante de surface (SU) des protéines d'enveloppe des virus des lymphomes / leucémies T chez les primates, regroupés sous la désignation PTLV, ces virus étant encore désignés HTLV chez l'homme et STLV chez le singe, à savoir de la région correspondant aux fragments protéiques délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des différentes souches des PTLV, ou de virus portant des séquences apparentées aux SU des PTLV,

pour la mise en œuvre de procédés de détection de toute souche de PTLV, à savoir de toute souche appartenant aux HTLV-1, HTLV-2, STLV-1, STLV-2, et STLV-3, ainsi que de toute souche de virus apparentés aux PTLV, à savoir de toute souche dont la séquence en acides aminés déduite de la séquence nucléotidique codant pour la région aminoterminal de la SU présente un taux d'homologie d'au moins environ 30% avec les séquences en acides aminés codées par les séquences nucléotidiques correspondantes chez les PTLV, notamment pour la détection de nouveaux variants des PTLV, ou de virus, nouveaux ou non, comportant des séquences apparentées aux SU des PTLV, le cas échéant dans de nouveaux contextes pathologiques, lesdits procédés comprenant une étape d'amplification, à partir d'un échantillon biologique susceptible de contenir des PTLV, et à l'aide des oligonucléotides 5' et 3' dégénérés susmentionnés utilisés en tant qu'amorces, du nombre de copies de fragments

nucléotidiques délimités en position 5' par l'oligonucléotide dégénéré en orientation 5', et en position 3' par l'oligonucléotide dégénéré en orientation 3', et une étape d'identification de la souche de PTLV contenue dans l'échantillon biologique à partir des fragments nucléotidiques amplifiés susmentionnés.

5 L'invention a plus particulièrement pour objet l'utilisation susmentionnée de couples d'oligonucléotides dégénérés tels que définis ci-dessus, caractérisée en ce que lesdits oligonucléotides sont choisis parmi ceux comprenant environ 15 à environ 30 nucléotides issus des séquences nucléotidiques codant pour des fragments protéiques délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un  
10 acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43, ou la souche NRA de HTLV-2 représentée par SEQ ID NO : 45, ou la souche de STLV-3 représentée par SEQ ID NO : 47, lesdits oligonucléotides dégénérés comprenant un mélange d'oligonucléotides issus de séquences codant pour une  
15 région déterminée d'environ 5 à 10 acides aminés des protéines d'enveloppe des différents souches de PTLV, et qui diffèrent entre eux par substitution d'au moins un nucléotide par un autre de manière à ce que chaque oligonucléotide soit susceptible de coder pour la région déterminée susmentionnée issue des fragments protéiques des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de  
20 HTLV-1, ou la souche NRA de HTLV-2, ou la souche de STLV-3 susmentionnées.

L'invention a également plus particulièrement pour objet l'utilisation susmentionnée de couples d'oligonucléotides dégénérés tels que définis ci-dessus, comprenant environ 15 à environ 30 nucléotides issus de séquences nucléotidiques codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus de fragments protéiques  
25 délimités par les acides aminés situés aux positions 80 à 245, et plus particulièrement aux positions 83 à 241, de la protéine d'enveloppe de la souche MT-2 de HTLV-1 (Gray et al., 1990, Virology, 177 : 391-395 ; n° d'accès Genbank M37747) représentée par SEQ ID NO : 43.

L'invention concerne plus particulièrement encore l'utilisation susmentionnée de  
30 couples d'oligonucléotides dégénérés tels que définis ci-dessus, issus de séquences nucléotidiques codant pour les fragments polypeptidiques 83-88, 140-145, 222-228, et 237-241, de la protéine d'enveloppe de la souche MT-2 de HTLV-1, à savoir les fragments suivants :



83-YL/VFPHW-88

140 – NFTQ/REV – 145

222-NYS/TCI/MVC-228

237-WHVLY-241

5 L'invention concerne plus particulièrement encore l'utilisation susmentionnée d'oligonucléotides dégénérés en orientation 5' issus du brin ADN (+) codant pour :

- le fragment polypeptidique 83-88 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (I) suivante :

TAYBTNTTYCCNCAYTGG (I) SEQ ID NO : 5

10 dans laquelle :

Y représente C ou T,

B représente C, G ou T,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

15 PTLVE5'83a TAYBTNTTYCCNCACTGG SEQ ID NO : 6

PTLVE5'83b TAYBTNTTYCCNCATTGG SEQ ID NO : 7

Y, B et N étant tels que définis ci-dessus,

- ou le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (II) suivante :

20 AAYTTYACNCARGARGT (II) SEQ ID NO : 8

dans laquelle :

Y représente C ou T,

R représente A ou G,

N représente A, C, G ou T,

25 telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

PTLVE5'140a AAYTTYACNCAAGAAGT SEQ ID NO : 9

PTLVE5'140b AAYTTYACNCAGGAAGT SEQ ID NO : 10

PTLVE5'140c AAYTTYACNCAAGAGGT SEQ ID NO : 11

PTLVE5'140d AAYTTYACNCAGGAGGT SEQ ID NO : 12

30 Y et N étant tels que définis ci-dessus.

L'invention concerne plus particulièrement encore l'utilisation susmentionnée, d'oligonucléotides dégénérés en orientation 3' issus du brin ADN (-) codant pour :

- le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (III) suivante :

NACYTCYTGNGTRAARTT (III) SEQ ID NO : 13

dans laquelle :

- 5 Y représente C ou T,  
R représente A ou G,  
N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

- |                |                    |                |
|----------------|--------------------|----------------|
| PTLVE3'145a    | NACYTCYTGNGTAAAATT | SEQ ID NO : 14 |
| 10 PTLVE3'145b | NACYTCYTGNGTGAAATT | SEQ ID NO : 15 |
| PTLVE3'145c    | NACYTCYTGNGTAAAGTT | SEQ ID NO : 16 |
| PTLVE3'145d    | NACYTCYTGNGTGAAGTT | SEQ ID NO : 17 |

Y et N étant tels que définis ci-dessus,

- 15 - ou le fragment polypeptidique 222-228 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (IV) suivante :

RMNACNATRCANSWRTARTT (IV) SEQ ID NO : 18

dans laquelle :

- 20 R représente A ou G,  
M représente A ou C,  
S représente C ou G,  
W représente A ou T,  
N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

- |                |                       |                |
|----------------|-----------------------|----------------|
| PTLVE3'228a    | RMNACNATRCANSAATAATT  | SEQ ID NO : 19 |
| 25 PTLVE3'228b | RMNACNATRCANSAGTAATT  | SEQ ID NO : 20 |
| PTLVE3'228c    | RMNACNATRCANSAATAGTT  | SEQ ID NO : 21 |
| PTLVE3'228d    | RMNACNATRCANSAGTAGTT  | SEQ ID NO : 22 |
| PTLVE3'228e    | RMNACNATRCANSTATAATT  | SEQ ID NO : 23 |
| PTLVE3'228f    | RMNACNATRCANSTGTAAATT | SEQ ID NO : 24 |
| 30 PTLVE3'228g | RMNACNATRCANSTATAGTT  | SEQ ID NO : 25 |
| PTLVE3'228h    | RMNACNATRCANSTGTAGTT  | SEQ ID NO : 26 |

R, M, S, et N étant tels que définis ci-dessus,

- ou le fragment polypeptidique 237-241 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (V) suivante :

RTANARNACRTGCCA (V) SEQ ID NO : 27

dans laquelle :

5 R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'241a RTANARNACATGCCA SEQ ID NO : 28

PTLVE3'241b RTANARNACGTGCCA SEQ ID NO : 29

10 R, et N étant tels que définis ci-dessus.

L'invention concerne également l'utilisation susmentionnées d'oligonucléotides tels que définis ci-dessus, comportant à leur extrémité 5' une séquence comprenant un site de restriction, tels que les sites EcoRI, de séquence GAATTC, ou BamHI, de séquence GGATCC.

15 A ce titre, l'invention concerne plus particulièrement l'utilisation susmentionnée d'oligonucléotides tel que définis ci-dessus, caractérisés en ce que les oligonucléotides 5' issus du brin ADN (+) correspondant aux fragments polypeptidiques 83-88 ou 140-145 comprennent en 5' une séquence GGAAGAATTC, et en ce que les oligonucléotides 3' issus du brin ADN (-) correspondant aux fragments polypeptidiques 140-145, 222-228, et 237-241  
20 comprennent en 5' une séquence GGAAGGATCC.

L'invention a également pour objet l'utilisation susmentionnée d'oligonucléotides tels que définis ci-dessus en tant que sondes, le cas échéant marquées, pour la mise en œuvre de procédés de détection susmentionnés de PTLV et de souches apparentées.

25 L'invention concerne également l'utilisation susmentionnée d'oligonucléotides tels que définis ci-dessus, en tant que couples d'amorces nucléotidiques pour la mise en œuvre de réactions de polymérisation en chaîne (PCR) pour la détection de toute souche de PTLV, ainsi que de toute souche de virus comportant des séquences apparentées aux SU des PTLV.

L'invention a plus particulièrement pour objet l'utilisation susmentionnée de couples d'amorces choisis de telle manière que :

30 - les oligonucléotides dégénérés 5' correspondent à un mélange d'oligonucléotides 5' issus d'une même région nucléotidique déterminée comprenant environ 15 à environ 30 nucléotides issus du brin ADN (+) et codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus de fragments protéiques délimités du côté N-terminal par un

acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 135 à 150 des protéines d'enveloppe des différentes souches des PTLV, notamment codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus du fragment protéique délimité du côté N-terminal par l'acide aminé situé à la position 83 et du côté C-terminal par l'acide aminé situé à la position 145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides 5' étant tels qu'ils diffèrent entre eux par substitution d'au moins un nucléotide par un autre de manière à ce que chaque oligonucléotide soit susceptible de coder pour la région déterminée susmentionnée issue des fragments protéiques des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de HTLV-1, ou la souche NRA de HTLV-2, ou la souche de STLTV-3 susmentionnées,

- les oligonucléotides dégénérés 3' correspondent à un mélange d'oligonucléotides 3' issus d'une même région nucléotidique déterminée comprenant environ 15 à environ 30 nucléotides issus du brin ADN (-) et codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus de fragments protéiques délimités du côté N-terminal par un acide aminé situé entre les positions 125 à 145, et du côté C-terminal par un acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des différentes souches des PTLV, notamment codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus du fragment protéique délimité du côté N-terminal par l'acide aminé situé à la position 140 et du côté C-terminal par l'acide aminé situé à la position 241 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides 3' étant tels qu'ils diffèrent entre eux par substitution d'au moins un nucléotide par un autre de manière à ce que chaque oligonucléotide soit susceptible de coder pour la région déterminée susmentionnée issue des fragments protéiques des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de HTLV-1, ou la souche NRA de HTLV-2, ou la souche de STLTV-3 susmentionnées,

étant bien entendu que lesdites amorces 5' et 3' susmentionnées ne peuvent pas être complémentaires l'une de l'autre.

L'invention concerne plus particulièrement l'utilisation susmentionnée de couples d'oligonucléotides dégénérés tels que définis ci-dessus, caractérisée en ce que les oligonucléotides dégénérés 5' sont choisis parmi les oligonucléotides 5' de formules (I) et (II) susmentionnées, et en ce que les oligonucléotides dégénérés 3' sont choisis parmi les oligonucléotides 3' de formules (III) à (V) susmentionnées.

L'invention concerne plus particulièrement l'utilisation susmentionnée de couples d'amorces tels que définis ci-dessus, caractérisée en ce que l'amorce 5' est choisie parmi les oligonucléotides 5' issus du brin ADN (+) correspondant aux fragments polypeptidiques 83 – 88 ou 140-145 définis ci-dessus, telles que les amorces PTLVE 5'83 a et b et PTLVE 5' 140 a à d susmentionnés, et en ce que l'amorce 3' est choisie parmi les oligonucléotides 3' issus du brin ADN (-) correspondant aux fragments polypeptidiques 140 – 145, 222 – 228 ou 237 – 241 définis ci-dessus, telles que les amorces PTLVE 3'145 a à d, PTLV3'228 a à h, et PTLVE 3 '145 a à d, PTLV 3'228 a à h, et PTLVE 3'241 a et b susmentionnées.

L'invention a plus particulièrement pour objet l'utilisation susmentionnée de couples d'oligonucléotides dégénérés tels que définis ci-dessus, lesdits oligonucléotides étant choisis de manière à ce qu'ils permettent l'amplification, à partir d'un échantillon biologique susceptible de contenir de l'ADN de PTLV, de séquences nucléotidiques codant pour des fragments protéiques comprenant une séquence délimitée du côté N-terminal par l'acide aminé situé en position 89, et du côté C-terminal par l'acide aminé situé en position 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43, ou comprenant une séquence analogue comprise dans la protéine d'enveloppe d'une autre souche de PTLV que HTLV-1, telle que la séquence délimitée du côté N-terminal par l'acide aminé situé en position 85, et du côté C-terminal par l'acide aminé situé en position 135 de la protéine d'enveloppe de la souche NRA de HTLV-2 représentée par SEQ ID NO : 45, ou la séquence délimitée du côté N-terminal par l'acide aminé situé en position 88, et du côté C-terminal par l'acide aminé situé en position 144 de la protéine d'enveloppe de la souche de STL-3 représentée par SEQ ID NO : 47.

L'invention a plus particulièrement pour objet encore l'utilisation susmentionnée de couples d'oligonucléotides dégénérés tels que définis ci-dessus, caractérisée en ce que les oligonucléotides dégénérés 5' sont choisis parmi les oligonucléotides 5' de formule (I) susmentionnée, et en ce que les oligonucléotides dégénérés 3' sont choisis parmi les oligonucléotides 3' de formules (III) à (V) susmentionnées.

L'invention concerne plus particulièrement l'utilisation susmentionnée de couples d'oligonucléotides dégénérés tels que définis ci-dessus, caractérisée en ce que :

- les oligonucléotides dégénérés 5' sont ceux de formule (I) suivante :

PTLVE5'83b TAYBTNTTYCCNCATTGG SEQ ID NO : 5

Y, B et N étant tels que définis ci-dessus,

- les oligonucléotides dégénérés 3' sont ceux de formule (III) suivante :

PTLVE3'145a NACYTCYTGNGTAAAATT

SEQ ID NO : 14

Y et N étant tels que définis ci-dessus.

L'invention concerne également les oligonucléotides tels que définis ci-dessus, en tant que tels.

5 A ce titre, l'invention a plus particulièrement pour objet les oligonucléotides tels que définis ci-dessus, correspondants :

- aux oligonucléotides dégénérés en orientation 5' issus du brin ADN (+) codant pour :

\* le fragment polypeptidique 83-88 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (I) suivante :

10 TAYBTNTTYCCNCAYTGG (I) SEQ ID NO : 5

dans laquelle :

Y représente C ou T,

B représente C, G ou T,

N représente A, C, G ou T,

15 telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

PTLVE5'83a TAYBTNTTYCCNCACTGG SEQ ID NO : 6

PTLVE5'83b TAYBTNTTYCCNCATTGG SEQ ID NO : 7

Y, B et N étant tels que définis ci-dessus,

20 \* le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (II) suivante :

AAYTTYACNCARGARGT (II) SEQ ID NO : 8

dans laquelle :

Y représente C ou T,

25 R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

PTLVE5'140a AAYTTYACNCAAGAAGT SEQ ID NO : 9

PTLVE5'140b AAYTTYACNCAGGAAGT SEQ ID NO : 10

30 PTLVE5'140c AAYTTYACNCAAGAGGT SEQ ID NO : 11

PTLVE5'140d AAYTTYACNCAGGAGGT SEQ ID NO : 12

Y et N étant tels que définis ci-dessus,

- aux oligonucléotides dégénérés en orientation 3' issus du brin ADN (-) codant pour :

\* le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (III) suivante :

5 NACYTCYTGNGTRAARTT (III) SEQ ID NO : 13

dans laquelle :

Y représente C ou T,

R représente A ou G,

N représente A, C, G ou T,

10 telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'145a NACYTCYTGNGTAAAATT SEQ ID NO : 14

PTLVE3'145b NACYTCYTGNGTGAAATT SEQ ID NO : 15

PTLVE3'145c NACYTCYTGNGTAAAGTT SEQ ID NO : 16

PTLVE3'145d NACYTCYTGNGTGAAGTT SEQ ID NO : 17

15 Y et N étant tels que définis ci-dessus,

\* le fragment polypeptidique 222-228 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (IV) suivante :

RMNACNATRCANSWRTARTT (IV) SEQ ID NO : 18

20 dans laquelle :

R représente A ou G,

M représente A ou C,

S représente C ou G,

W représente A ou T,

25 N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'228a RMNACNATRCANSAATAATT SEQ ID NO : 19

PTLVE3'228b RMNACNATRCANSAGTAATT SEQ ID NO : 20

PTLVE3'228c RMNACNATRCANSAATAGTT SEQ ID NO : 21

30 PTLVE3'228d RMNACNATRCANSAGTAGTT SEQ ID NO : 22

PTLVE3'228e RMNACNATRCANSTATAATT SEQ ID NO : 23

PTLVE3'228f RMNACNATRCANSTGTAATT SEQ ID NO : 24

PTLVE3'228g RMNACNATRCANSTATAGTT SEQ ID NO : 25

PTLVE3'228h RMNACNATRCANSTGTAGTT SEQ ID NO : 26

R, M, S, et N étant tels que définis ci-dessus,

\* le fragment polypeptidique 237-241 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (V) suivante :

RTANARNACRTGCCA (V) SEQ ID NO : 27

dans laquelle :

R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'241a RTANARNACATGCCA SEQ ID NO : 28

PTLVE3'241b RTANARNACGTGCCA SEQ ID NO : 29

R, et N étant tels que définis ci-dessus.

L'invention a également pour objet un procédé de détection de toute souche de PTLV, à savoir de toute souche appartenant aux HTLV-1, HTLV-2, STLV-1, STLV-2, et STLV-3, ainsi que de toute souche de virus comportant des séquences apparentées aux SU des PTLV, telles que définies ci-dessus, caractérisé en ce qu'il comprend :

- la mise en contact d'un couple d'oligonucléotides dégénérés 5' et 3' tels que définis ci-dessus, avec l'ADN génomique ou l'ADN complémentaire dérivé à partir d'ARN extraits du contenu d'un échantillon biologique (tels que cellules sanguines, de moelle osseuse, biopsies, notamment de peau ou autres organes, ou frottis) susceptible de contenir des PTLV tels que définis ci-dessus,

- l'amplification de fragments d'ADN codant pour un fragment des protéines d'enveloppe des différentes souches des PTLV tel que défini ci-dessus,

- la détection des fragments d'ADN amplifiés lors de l'étape précédente, cette détection pouvant être corrélée à la détection et le cas échéant à l'identification de PTLV tels que définis ci-dessus dans ledit échantillon biologique.

L'invention concerne également un procédé de détection de toute souche de PTLV tel que défini ci-dessus, caractérisé en ce que l'étape d'amplification comprend la mise en œuvre de deux réactions d'amplification, la deuxième réaction étant effectuée sur un échantillon de produits obtenus dans le cadre de la première réaction à l'aide des mêmes oligonucléotides 5' que dans le cas de la première réaction, et d'oligonucléotides 3' différents de ceux utilisés dans la première réaction, à savoir des amorces 3' dites « nichées » hybridant avec une région



située plus en amont de la séquence codant pour la SU que les amorces 3' utilisées dans la première réaction.

L'invention a également pour objet un procédé de détection de toute souche de PTLV tel que défini ci-dessus, caractérisé en ce qu'il comprend :

- 5 - une première réaction d'amplification génique effectuée à l'aide de couples d'oligonucléotides dégénérés choisis parmi les couples :
- \* oligonucléotides de formule (I)/oligonucléotides de formule (IV), ou
  - \* oligonucléotides de formule (I)/oligonucléotides de formule (V), ou
  - \* oligonucléotides de formule (II)/oligonucléotides de formule (V),
- 10 - et une deuxième étape d'amplification du nombre de copies de fragments d'ADN obtenus lors de l'étape précédente à l'aide de couples d'oligonucléotides dégénérés choisis respectivement parmi les couples :
- \* oligonucléotides de formule (I)/oligonucléotides de formule (III), ou
  - \* oligonucléotides de formule (I)/oligonucléotides de formule (III ou IV), ou
  - 15 \* oligonucléotides de formule (II)/oligonucléotides de formule (IV),
- la détection des fragments d'ADN amplifiés lors de l'étape précédente, cette détection pouvant être corrélée à la détection et le cas échéant à l'identification de PTLV dans l'échantillon biologique.

L'invention concerne également un procédé de détection de toute souche de PTLV tel  
20 que défini ci-dessus, caractérisé en ce qu'il comprend :

- une première réaction d'amplification génique effectuée à l'aide de couples  
d'oligonucléotides dégénérés choisis de telle manière que :
- \* les oligonucléotides dégénérés 5' sont ceux de formule (I) suivante :
- PTLVE5'83b    TAYBTNTTYCCNCATTGG                      SEQ ID NO : 5
- Y, B et N étant tels que définis ci-dessus,
- \* les oligonucléotides dégénérés 3' sont ceux de formule (V) suivante :
- PTLVE3'241b    RTANARNACGTGCCA                      SEQ ID NO : 29
- R, et N étant tels que définis ci-dessus,
- une deuxième réaction d'amplification génique effectuée à l'aide de couples  
d'oligonucléotides dégénérés choisis de telle manière que :
- \* les oligonucléotides dégénérés 5' sont ceux de formule (I) suivante :
- PTLVE5'83b    TAYBTNTTYCCNCATTGG                      SEQ ID NO : 5
- Y, B et N étant tels que définis ci-dessus.

\* les oligonucléotides dégénérés 3' sont ceux de formule (III) suivante :

PTLVE3'145a      NACYTCYTGNGTAAAATT

SEQ ID NO : 14

Y et N étant tels que définis ci-dessus.

L'invention concerne plus particulièrement un procédé de détection tel que défini ci-dessus, caractérisé en ce que l'étape d'amplification est effectuée dans les conditions suivantes :

– dénaturation à 94°C pendant 5 min,

– une première réaction PCR en conditions dites de « touch down » effectuée dans un milieu contenant de la Taq polymérase ou autres ADN polymérases fonctionnant à haute température, cette première réaction PCR comprenant :

. 15 cycles d'affilée « touch down » variant par la température d'élongation qui diminue de 1°C à chaque cycle comprenant :

\* une étape de dénaturation à 94°C pendant 15 sec,

\* une étape combinant appariement et élongation à une température variant entre 65°C et 50 °C pendant 20 sec,

. 30 cycles classiques comprenant :

\* une étape de dénaturation à 94°C pendant 15 sec,

\* une étape d'appariement à 50°C pendant 30 sec,

\* une étape d'élongation à 72°C pendant 30 sec,

– une deuxième réaction PCR effectuée sur un échantillon de produits obtenus dans le cadre de la première réaction PCR susmentionnée à l'aide de la même amorce 5' que dans le cas de la réaction PCR précédente, et d'une amorce 3' différente de celle utilisée dans la réaction PCR précédente, à savoir une amorce 3' dite « nichée » hybridant avec une région située plus en amont de la séquence codant pour la SU que l'amorce 3' utilisée à l'étape précédente.

L'invention a plus particulièrement pour objet un procédé de détection tel que défini ci-dessus, caractérisé en ce que l'étape de détection, et le cas échéant d'identification, est effectuée dans les conditions suivantes :

- ligation directe des fragments amplifiés lors de l'étape d'amplification dans un plasmide tel que pCR4-TOPO (Invitrogen),

- transformation de bactéries avec le plasmide susmentionné comprenant un gène marqueur tel qu'un gène de résistance à un antibiotique, notamment à la kanamycine,

- repiquage de colonies bactériennes (notamment entre 10 et 100), culture, extraction de l'ADN, et séquençage (notamment à l'aide des amorces universelles T3 ou T7 dans le cas de l'utilisation du vecteur pCR4-TOPO).

5 L'invention concerne également une trousse ou kit pour la mise en œuvre d'un procédé de détection tel que défini ci-dessus, caractérisé en ce qu'il comprend un couple d'oligonucléotides dégénérés susmentionnés, et, le cas échéant, les réactifs nécessaires à la mise en œuvre de la réaction d'amplification par PCR et pour la détection des fragments amplifiés.

10 L'invention a également pour objet l'application du procédé de détection défini ci-dessus au diagnostic de pathologies liées à une infection par un PTLV, ou par un virus comportant des séquences apparentées aux SU des PTLV, chez l'homme ou l'animal, telles que les hémopathies, les maladies auto-immunes, les maladies inflammatoires, les maladies dégénératives.

15 A ce titre, l'invention concerne toute méthode de diagnostic in vitro de pathologies susmentionnées par mise en œuvre d'un procédé de détection défini ci-dessus, la détection de fragments d'ADN amplifiés pouvant être corrélée au diagnostic desdites pathologies.

Le cas échéant, les méthodes de diagnostic in vitro de l'invention comprenant une étape supplémentaire d'identification de PTLV ou virus apparentés aux PTLV présents dans l'échantillon biologique, par séquençage des fragments d'ADN amplifiés.

20 L'invention a également pour objet l'application du procédé de détection défini ci-dessus, au criblage et à l'identification de nouveaux agents infectieux chez l'homme ou l'animal, et plus particulièrement de nouvelles souches (ou variants) de virus susceptibles d'être classés dans les PTLV, ou de virus comportant des séquences apparentées aux SU des PTLV.

25 Les méthodes de criblage et d'identification susmentionnées de nouveaux agents infectieux sont effectuées par mise en œuvre d'un procédé de détection défini ci-dessus et comprennent une étape supplémentaire d'identification des nouveaux variants de PTLV ou de virus apparentés par séquençage des fragments d'ADN amplifiés.

30 L'invention concerne également l'application du procédé de détection défini ci-dessus au criblage de gènes de prédisposition ou de résistance à des pathologies chez l'homme ou l'animal liées à la présence de séquences des PTLV ou de séquences apparentées, ou à une infection par un PTLV, telles que les hémopathies, les maladies auto-immunes, les maladies dégénératives.

L'invention a également pour objet l'application du procédé de détection défini ci-dessus, au criblage ou à la conception de nouveaux agents thérapeutiques comprenant des séquences entières ou partielles des protéines d'enveloppe de nouveaux variants de PTLV ainsi détectés.

5 L'invention concerne également l'application du procédé de détection tel que défini ci-dessus, au criblage ou à la conception de nouveaux vecteurs de thérapie cellulaire utilisant les propriétés de tropisme des séquences entières ou partielles des protéines d'enveloppe de nouveaux variants de PTLV ainsi détectés.

10 L'invention a également pour objet les variants de type HTLV-1 tels qu'obtenus par mise en œuvre d'un procédé de détection défini ci-dessus, correspondants :

- au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 31 suivante :

15 I K K P N P N G G G Y Y L A S Y S D  
P C S L K C P Y L G C Q S W T C P Y  
T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu arginine (R) en position 94, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu proline (P) et un résidu leucine (L) indiqués en gras et  
20 soulignés,

et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 30 suivante :

25 ATT AAA AAG CCA AAC CCA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 281, C en position 302, et G en position 333, sont respectivement remplacés par C, T, et A indiqués en gras et soulignés.

30 - au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 33 suivante :

V K K P N R N G G G Y Y L A S Y S D  
P C S L K C P Y L G C Q S W T C P Y  
T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu isoleucine (I) en position 89, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu valine (V) et un résidu leucine (L) indiqués en gras et soulignés,

et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 32 suivante :

GTT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle A en position 266, C en position 302, et G en position 333, sont respectivement remplacés par G, T, et A indiqués en gras et soulignés.

- au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 35 suivante :

I	K	K	P	N	R	N	G	G	G	Y	Y	<u>L</u>	A	S	Y	S	D
P	C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y
T	G	A	V	S	S	P	Y	W	K	F	Q	Q	D	V			

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu sérine (S) en position 101, est remplacé par un résidu leucine (L) indiqué en gras et souligné,

et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 34 suivante :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAA CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle C en position 302, G en position 333, et G en position 408, sont respectivement remplacés par T, A, et A indiqués en gras et soulignés.

- au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 37 suivante :

I	K	K	P	N	R	N	G	G	G	Y	Y	L	A	S	Y	S	D
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

P C S L K C P Y L G C Q S W T C P Y  
T G P V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu alanine (A) en position 127, est remplacé par un résidu proline (P) indiqué en gras et souligné,

et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 36 suivante :

GTT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
10 CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA CCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 379, est remplacé par C indiqué en gras et souligné,

15 - au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 39 suivante :

I K K P N R N G G G Y H S A S Y S D P  
C S L K C P Y L G C Q S W T C P Y A G  
A V S S P Y W K F Q Q D V N F T Q E V

20 à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu tyrosine (Y) en position 100, et le résidu thréonine (T) en position 125, sont respectivement remplacés par un résidu histidine (H) et un résidu alanine (A) indiqués en gras et soulignés,

25 et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 38 suivante :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT CAT TCA GCC TCT TAT TCA  
GAC CCT TGT TCC TTA AAG TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC  
CCC TAT GCA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC  
30 AAT TTT ACC CAG GAA GTA

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 435 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle T en position 298, A en position 373, T en position 426, A en position 429, et T en position 435, sont respectivement remplacés par C, G, C, G, et A  
35 indiqués en gras et soulignés.

L'invention a également pour objet le variant de type HTLV-2 tel qu'obtenu par mise en œuvre d'un procédé de détection défini ci-dessus, caractérisé en ce que:

- sa protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO :

41 suivante :

5 I R K P N R Q G L G Y Y S P S Y N D  
P C S L Q C P Y L G S Q S W T C P Y  
T A P V S T P S W N F H S D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 85 à 135 de la protéine d'enveloppe de la souche prototype NRA de HTLV-2 (décrite par Lee et al., 1993. Virology 196, 57-69 ; n° d'accès Genbank L20734.1), dans laquelle les résidus lysine (K) en position 86, cystéine (C) en position 113, glycine (G) en position 122, sérine (S) en position 126, et lysine (K) en position 130, sont respectivement remplacés par les résidus arginine (R), sérine (S), alanine (A), thréonine (T), et asparagine (N) indiqués en gras et soulignés,

15 - la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 40 suivante :

ATA AGA AAG CCA AAC AGA CAG GGC CTA GGG TAC TAC TCG CCT TCC TAC AAT GAC  
CCT TGC TCG CTA CAA TGC CCC TAC TTG GGC TCC CAA TCA TGG ACA TGC CCA TAC  
ACG GCC CCC GTC TCC ACT CCA TCC TGG AAT TTT CAT TCA GAT GTA

20 à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 253 à 405 de la séquence codant pour la protéine d'enveloppe de la souche NRA de HTLV-2, dans laquelle A en position 257, G en position 258, T en position 267, A en position 282, C en position 294, T en position 300, A en position 333, G en position 338, G en position 365, G en position 377, G en position 390, et C en position 396, sont respectivement remplacés par G, A, C, G, T, C, G, C, C, C, T, et T indiqués en gras et soulignés.

25 L'invention concerne également les polypeptides délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43, ou de la souche NRA de HTLV-2 représentée par SEQ ID NO : 45, ou de la souche de STL-3 représentée par SEQ ID NO : 47, ou de virus portant des séquences apparentées aux SU des PTLV, ou délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 135 à 150 desdites protéines d'enveloppe des différentes souches des PTLV.

L'invention a également pour objet les polypeptides définis ci-dessus, choisis parmi :

- le polypeptide délimité du côté N-terminal par l'acide aminé situé à la position 83 ou 89, et du côté C-terminal par l'acide aminé situé à la position 139 ou 145, de la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43,

5 - le polypeptide délimité du côté N-terminal par l'acide aminé situé à la position 79 ou 85, et du côté C-terminal par l'acide aminé situé à la position 135 ou 141, de la protéine d'enveloppe de la souche NRA de HTLV-2 représentée par SEQ ID NO : 45,

- le polypeptide délimité du côté N-terminal par l'acide aminé situé à la position 82 ou 88, et du côté C-terminal par l'acide aminé situé à la position 138 ou 144, de la protéine  
10 d'enveloppe de la souche de STLV-3 représentée par SEQ ID NO : 47.

L'invention concerne également les polypeptides codés par les fragments d'ADN amplifiés dans le cadre du procédé de détection défini ci-dessus, des variants de type HTLV-1 à HTLV-2 susmentionnés, caractérisés en ce qu'ils comprennent les séquences peptidiques suivantes :

15 - polypeptide 1 (SEQ ID NO : 31) :

I	K	K	P	N	<u>P</u>	N	G	G	G	Y	Y	<u>L</u>	A	S	Y	S	D
P	C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y
T	G	A	V	S	S	P	Y	W	K	F	Q	Q	D	V			

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés  
20 aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu arginine (R) en position 94, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu proline (P) et un résidu leucine (L) indiqués en gras et soulignés,

-polypeptide 2 (SEQ ID NO : 33):

25	<u>V</u>	K	K	P	N	R	N	G	G	G	Y	Y	<u>L</u>	A	S	Y	S	D
	P	C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y
	T	G	A	V	S	S	P	Y	W	K	F	Q	Q	D	V			

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés  
30 aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu isoleucine (I) en position 89, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu valine (V) et un résidu leucine (L) indiqués en gras et soulignés,

- polypeptide 3 (SEQ ID NO : 35):

	I	K	K	P	N	R	N	G	G	G	Y	Y	<u>L</u>	A	S	Y	S	D
35	P	C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y



T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu sérine (S) en position 101, est remplacé par un résidu leucine (L) indiqué en gras et souligné.

- polypeptide 4 (SEQ ID NO : 37):

I K K P N R N G G G Y Y L A S Y S D  
P C S L K C P Y L G C Q S W T C P Y  
T G P V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu alanine (A) en position 127, est remplacé par un résidu proline (P) indiqué en gras et souligné,

- polypeptide 5 (SEQ ID NO : 39) :

I K K P N R N G G G Y H S A S Y S D P  
C S L K C P Y L G C Q S W T C P Y A G  
A V S S P Y W K F Q Q D V N F T Q E V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu tyrosine (Y) en position 100, et le résidu thréonine (T) en position 125, sont respectivement remplacés par un résidu histidine (H) et un résidu alanine (A) indiqués en gras et soulignés,

- polypeptide 6 (SEQ ID NO : 41) :

I R K P N R Q G L G Y Y S P S Y N D  
P C S L Q C P Y L G S Q S W T C P Y  
T A P V S T P S W N F H S D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 85 à 135 de la protéine d'enveloppe de la souche prototype NRA de HTLV-2, dans laquelle les résidus lysine (K) en position 86, cystéine (C) en position 113, glycine (G) en position 122, sérine (S) en position 126, et lysine (K) en position 130, sont respectivement remplacés par les résidus arginine (R), sérine (S), alanine (A), thréonine (T), et asparagine (N) indiqués en gras et soulignés.

L'invention a également pour objet les acides nucléiques caractérisés en ce qu'ils codent pour un polypeptide tel que défini ci-dessus.

L'invention concerne plus précisément les acides nucléiques susmentionnés, comprenant les séquences nucléotidiques suivantes :

- acide nucléique 1 a (SEQ ID NO : 30):

ATT AAA AAG CCA AAC CCA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
 5 CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
 ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 281, C en position 302, et G en position 333, sont  
 10 respectivement remplacés par C, T, et A indiqués en gras et soulignés,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 1 susmentionné,

- acide nucléique 2 a (SEQ ID NO : 32):

GTT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
 15 CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
 ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle A en position 266, C en position 302, et G en position 333, sont  
 20 respectivement remplacés par G, T, et A indiqués en gras et soulignés,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 2 susmentionné,

- acide nucléique 3 a (SEQ ID NO : 34) :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
 25 CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
 ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAA CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle C en position 302, G en position 333, et G en position 408, sont  
 30 respectivement remplacés par T, A, et A indiqués en gras et soulignés,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 3 la revendication 24,

- acide nucléique 4 a (SEQ ID NO : 36):

GTT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
 35 CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT

ACA GGA CCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 379, est remplacé par C indiqué en gras et souligné,

5 ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 4 susmentionné,

- acide nucléique 5 a (SEQ ID NO : 38) :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT CAT TCA GCC TCT TAT TCA  
 GAC CCT TGT TCC TTA AAG TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC  
 10 CCC TAT GCA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC  
 AAT TTT ACC CAG GAA GTA

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 435 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle T en position 298, A en position 373, T en position 426, A en position 429, et T en position 435, sont respectivement remplacés par C, G, C, G, et A  
 15 indiqués en gras et soulignés.

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 5 susmentionné,

- acide nucléique 6 a (SEQ ID NO : 40):

20 ATA AGA AAG CCA AAC AGA CAG GGC CTA GGG TAC TAC TCG CCT TCC TAC AAT GAC  
 CCT TGC TCG CTA CAA TGC CCC TAC TTG GGC TCC CAA TCA TGG ACA TGC CCA TAC  
 ACG GCC CCC GTC TCC ACT CCA TCC TGG AAT TTT CAT TCA GAT GTA

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 253 à 405 de la séquence codant pour la protéine d'enveloppe de la souche NRA de HTLV-2, dans laquelle A en position 257, G en position 258, T en position 267, A en position 282, C en position 294, T en position 300, A en position 333, G en position 338, G en position 365, G en position 377, G en position 390, et C en position 396, sont respectivement remplacés par G, A, C, G, T, C, G, C, C, C, T, et T indiqués en gras et soulignés,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant  
 30 pour le polypeptide 6 susmentionné.

L'invention concerne également les anticorps polyclonaux ou monoclonaux dirigés contre un nouveau variant de type HTLV – 1 ou HTLV – 2 tel que défini ci-dessus, ou contre un polypeptide défini ci-dessus, lesdits anticorps étant tels qu'obtenus par immunisation d'un animal approprié avec un polypeptide susmentionné.

L'invention a également pour objet toute composition pharmaceutique, notamment vaccins ou vecteurs thérapeutiques, conçus à partir des nouveaux variants de type HTLV-1 ou HTLV-2 tels que définis ci-dessus, et plus particulièrement toute composition pharmaceutique comprenant un polypeptide selon l'invention tel que défini ci-dessus, notamment les polypeptides 1 à 6 définis ci-dessus, ou un acide nucléique 1a à 6a défini ci-dessus, ou des anticorps susmentionnés, le cas échéant en association avec un véhicule pharmaceutiquement acceptable.

L'invention concerne également l'utilisation des nouveaux variants de type HTLV-1 ou HTLV-2 tels que définis ci-dessus, ou des polypeptides selon l'invention tels que définis ci-dessus, notamment des polypeptides 1 à 6, ou des acides nucléiques 1a à 6a définis ci-dessus, ou des anticorps susmentionnés, pour la préparation de médicaments destinés à la prévention ou au traitement des infections d'un individu par les PTLV susmentionnés, ainsi que des pathologies définies ci-dessus liées à l'infection par ces PTLV.

L'invention sera davantage illustrée à l'aide de la description détaillée qui suit de l'obtention d'amorces selon l'invention et de leur utilisation pour la détection de nouveaux variants de HTLV.

## **I. MISE AU POINT D'OUTILS MOLECULAIRES ET DE STRATEGIES DE DETECTION DE SEQUENCES PAN-PTLV PAR AMPLIFICATION, CLONAGE ET SEQUENÇAGE, DE SEQUENCES NUCLEOTIDIQUES APPARENTEES AUX SU DES ENVELOPPES PTLV.**

### **1. Recherche de motifs peptidiques conservés en N-terminus des SU de PTLV**

La question principale résolue par les inventeurs est la mise au point d'outils et d'une méthode permettant d'amplifier, de cloner et d'identifier toute séquence nucléotidique apparentée à la SU des PTLV qui est responsable de la reconnaissance de leur récepteur cellulaire (Kim et coll, 2000). Pour cela, les inventeurs ont cherché des motifs peptidiques conservés dans la SU des enveloppes de PTLV pour en déduire des séquences nucléotidiques permettant de les représenter toutes. Ces motifs peptidiques devaient répondre idéalement aux 5 critères suivants, selon leur importance décroissante:

- Etre conservés parmi la plupart, sinon la totalité, des séquences déjà décrites d'enveloppes de PTLV. Une telle conservation serait une garantie de leur efficacité potentielle dans la détection de nouvelles séquences de type PTLV.

- Représenter au moins 5 acides aminés conservés de la SU des PTLV, afin d'en dériver une séquence minimale de 15 nucléotides. En effet, étant donnée la complexité des génomes eucaryotes, ce minimum de 15 nucléotides est requis pour permettre la détection spécifique d'une séquence nucléotidique donnée.

5 - Permettre l'amplification de séquences situées en amont du motif C I/M V C qui est conservé dans la SU des PTLV et semble analogue au motif CWLC décrit dans la SU des MuLV (Sitbon et coll, 1991). Ce motif semble en effet, marquer une région charnière entre, en son amont, la partie de la SU responsable de la reconnaissance du récepteur et, en son aval, les domaines carboxy terminaux de la SU impliqués dans l'association avec la TM et des étapes  
10 de l'entrée virale postérieures à la reconnaissance du récepteur (Battini, et coll, 1992; Battini et coll, 1995; Lavillette et coll, 1998; Kim et coll, 2000; Lavillette et coll, 2001). Etre suffisamment distants les uns des autres pour permettre l'amplification d'un fragment dont la longueur augmenterait les chances de détection d'un polymorphisme éventuel entre différentes séquences.

15 - Etre situés de façon à permettre deux réactions d'amplifications d'ADN successives, dont la deuxième, nichée, est réalisée à partir des produits de la première amplification, et permettrait l'amplification d'un fragment interne au premier fragment amplifié. Cette amplification nichée permettrait d'augmenter la probabilité d'amplification d'un fragment qui corresponde bien à une séquence apparentée à la SU des PTLV.

20 Selon ces critères, les inventeurs ont identifié les motifs d'acides aminés suivants, présents dans toutes ou la quasi-totalité des SU connues de PTLV, et pouvant se prêter au développement de cette stratégie :

- Motif peptidique 1 : Y L/V F P H W
- 25 - Motif peptidique 2 : N F T Q/R E V
- Motif peptidique 3 : N Y S/T C I/M V C
- Motif peptidique 4 : W H V L Y

## 30 **2. Oligonucléotides dégénérés de synthèse correspondant aux motifs conservés dans la partie amino terminale de la SU des PTLV**

À partir des séquences d'acides aminés des motifs peptidiques conservés identifiés ci-dessus et suivant la correspondance nucléotidique en application du code génétique eucaryote, les inventeurs ont déterminé des séquences nucléotidiques dégénérées (SND) qui ont servi de

base pour la conception d'oligonucléotides de synthèse (OS). Plusieurs critères ont présidé à la conception d'OS correspondants à ces SND :

- Lorsque la multiplication des positions dégénérées dans une SND faisait que la complexité de l'OS correspondant dépasse 512 oligonucléotides dans le mélange de synthèse, la synthèse d'OS supplémentaires pour cette SND est alors effectuée pour lever une partie de cette complexité.

- La synthèse d'un ou 2 OS supplémentaires, dans la limite de 4 OS par SND, est effectuée même pour des complexités inférieures à 512, lorsque ces OS supplémentaires lèvent significativement la complexité de l'OS dégénéré initial.

- Les séquences des OS 5', dont l'élongation par les polymérases d'ADN doit correspondre aux acides aminés situés en aval du motif peptidique considéré (motifs 1 et 2), sont celles du brin ADN (+), alors que celles des OS 3', à partir desquels l'élongation doit correspondre aux acides aminés situés en amont du motif peptidique considéré (motifs 2, 3 et 4), sont celles du brin ADN (-). Les OS correspondants au motif peptidique 2 ont été synthétisés sur les deux brins, pour pouvoir effectuer une élongation dans les deux sens.

- Chaque OS comprend des nucléotides supplémentaires permettant d'introduire en 5' la séquence correspondant à un site de restriction, EcoRI pour les OS 5', BamHI pour les OS 3', et dans tous les cas une séquence GGAA 5'-terminale favorisant l'arrimage des polymérases et nucléases en amont du site de restriction.

- Suivant ces critères, les OS PTLVE5' (83 a et b, 140 a à d) et PTLVE3' (145 a à d, 228 a à h, 241 a et b) (pour *Primate T-Leukemia Virus-like Env*), définis ci-dessus ont été synthétisés respectivement pour des élongations en 5' ou 3' du motif ciblé.

### 3. Mise au point des conditions d'amplification avec les oligonucléotidiques sur des séquences témoins

Pour la mise au point de l'amplification de séquences reconnues par les OS décrits ci-dessus, les inventeurs ont utilisé des préparations ADN témoins de plasmide contenant la séquence d'enveloppe HTLV-1 et des préparations témoins dépourvues de cette séquence. La stratégie d'amplification d'ADN qui a été sélectionnée consiste à enchaîner deux réactions d'amplification par un mélange des polymérases Taq et Pwo sur thermocycler dans les conditions dites de "touch-down" et combinant 2 couples d'OS différents.

Les premiers résultats d'amplification probants et reproductibles (amplification spécifiques de séquences HTLV sans amplification sur les préparations témoins) sont ceux obtenus avec la combinaison des OS PTLVE5'83b et PTLVE3'240b, pour la première réaction d'amplification, suivie d'une 2<sup>ème</sup> réaction combinant les OS PTLVE5'83b et PTLVE3'146a sur un échantillon de la 1<sup>ère</sup> réaction. Dans les deux cas les conditions de "touch-down" incluent 15 cycles combinant chacun dénaturation à 94°C suivi d'une étape d'appariement et d'élongation effectuée à la même température, cette température étant comprise pour chaque cycle entre 65 et 50°C avec un pas décroissant de 1°C entre le 1<sup>er</sup> et le 15<sup>e</sup> cycle. Ces 15 cycles sont suivis de 30 cycles classiques d'amplification avec une température d'appariement à 50°C et d'élongation à 72°C.

#### 4. Construction et séquençage d'une banque de fragments amplifiés à partir des réactions d'amplification

Un échantillon de la 2<sup>ème</sup> réaction d'amplification décrite ci-dessus est utilisé pour générer une banque des séquences amplifiées. Pour cela 4 µl sur les 50 µl de la 2<sup>ème</sup> réaction est utilisé pour ligation dans un vecteur de type pCR4-TOPO (Invitrogen) et transformation de bactéries. Entre 10 et 100 colonies résistantes à la kanamycine sont repiquées pour chaque ligation et mises en culture. L'ADN plasmidique de chaque colonie est analysé par séquençage en utilisant des séquences amorces universelles T3 et T7 du vecteur.

## II PREMIERS RESULTATS OBTENUS A PARTIR D'ECHANTILLONS HUMAINS ET DE PRIMATES

Les conditions décrites ci-dessus ont été appliquées à trois types d'échantillons d'ADN :

- Des échantillons d'ADN génomique de "patients séroindéterminés", caractérisés par une sérologie suggérant une infection antérieure par HTLV mais chez lesquels aucun diagnostic définitif n'a pu être établi. Chez ces patients, notamment, une recherche par amplification d'ADN de séquences HTLV *gag*, *pol* ou *tax* sont négatives.

- Des échantillons d'ADN génomique de "patients HTLV-1" chez lesquels une infection HTLV-1 caractéristique a été identifiée.

- Des échantillons d'ADN génomique de singes Mangabey agiles (*Cercocebus Agilis*) qui présentent une sérologie PTLV positive et chez lesquels des séquences Tax HTLV-1 ou STLV-L ont pu être amplifiées.

L'application de la méthode décrite ci-dessus a permis de détecter la présence de séquences de type SU de PTLV chez les trois types d'échantillons, y compris chez les "patients séroindéterminés".

5 L'analyse des séquences et de leurs capacités codantes au niveau de la région de SU concernée a permis de faire les observations suivantes:

### **1. Résultats obtenus sur "patients séroindéterminés"**

10 En appliquant la méthode décrite ci-dessus sur l'ADN d'un "patient séroindéterminé" (échantillon No. 424), décrit comme ne portant pas de séquence de type HTLV, les inventeurs ont pourtant pu amplifier et caractériser des séquences de type SU de PTLV.

Au niveau nucléotidique, les séquences identifiées à partir de l'échantillon No. 424 sont de plusieurs types : des séquences HTLV-1 identiques à celles déjà décrites dans la littérature et de nouveaux variants. Au niveau codant, les séquences nucléotidiques se traduisent en trois types de séquences :

- 15
- Des séquences d'acides aminés identiques à celles des souches HTLV-1 déjà connues.
  - Des variants de souches HTLV-1 avec 1 ou 2 résidus jamais décrits auparavant.
  - Des variants de souches HTLV-1 avec 1 ou 2 résidus décrits comme communs aux seules souches HTLV-2 ou STL-V-L.

### **20 2. Résultats obtenus sur "patients HTLV-1 typiques"**

Au niveau nucléotidique, les séquences amplifiées à partir de l'échantillon provenant du "patient HTLV-1" (échantillon No.422) sont soit typiquement HTLV-1, telles que déjà décrites dans la littérature, soit des variants avec des répercussions sur la capacité codante. Au niveau codant, les séquences nucléotidiques se traduisent en trois types de séquences :

- 25
- Des séquences d'acides aminés identiques aux souches connues HTLV-1.
  - Des variants HTLV-1 avec 1 ou 2 résidus typiques des HTLV-2, combinés ou pas avec des résidus jamais décrits auparavant.
  - Des variants HTLV-2 combinant quelques résidus décrits pour être communs aux seules souches HTLV-2 ou STL-V-L, ceci combinés ou pas avec des résidus jamais décrits
- 30 auparavant.



### 3. Résultats obtenus sur singes *Cercocebus Agilis*

La méthode de l'invention a aussi permis d'amplifier des séquences de type SU de PTLV chez tous les Mangabey Agiles (*Cercocebus Agilis*) testés qui avaient été identifiés comme séropositifs pour PTLV. Au niveau nucléotidique, les séquences amplifiées à partir de ces singes sont soit celles des isolats déjà décrits auparavant, soit des variants nucléotidiques avec répercussions sur la capacité codante. Au niveau codant, les séquences nucléotidiques se traduisent en trois types de séquences :

- Des séquences d'acides aminés identiques aux souches connues HTLV-1.
- Des séquences d'acides aminés identiques à l'isolat STLV-3/CTO-604 tout récemment décrit chez un Mangabey à tête rouge (*Cercocebus Torquatus*) (Meertens et coll, 2002)
- Des séquences d'acides aminés du type STLV-3/CTO-604 avec 1 ou 2 résidus typiques des HTLV-2

### III. BIBLIOGRAPHIE

1. Battini, J. L., O. Danos, and J. M. Heard. 1995. Receptor-binding domain of murine leukemia virus envelope glycoproteins. J Virol. 69(2):713-719.
2. Battini, J. L., J. M. Heard, and O. Danos. 1992. Receptor choice determinants in the envelope glycoproteins of amphotropic, xenotropic, and polytropic murine leukemia viruses. J Virol. 66(3):1468-75.
3. Kim, F. J., I. Seiliez, C. Denesvre, D. Lavillette, F. L. Cosset, and M. Sitbon. 2000. Definition of an amino-terminal domain of the human T-cell leukemia virus type 1 envelope surface unit that extends the fusogenic range of an ecotropic murine leukemia virus. J Biol Chem. 275(31):23417-20.
4. Lavillette, D., M. Maurice, C. Roche, S. J. Russell, M. Sitbon, and F. L. Cosset. 1998. A proline-rich motif downstream of the receptor binding domain modulates conformation and fusogenicity of murine retroviral envelopes. J Virol. 72(12):9955-65.
5. Lavillette, D., A. Ruggieri, S. J. Russell, and F. L. Cosset. 2000. Activation of a cell entry pathway common to type C mammalian retroviruses by soluble envelope fragments. J Virol. 74(1):295-304.
6. Meertens, L., R. Mahieux, P. Mauclore, J. Lewis, and A. Gessain. 2002. Complete Sequence of a Novel Highly Divergent Simian T-Cell Lymphotropic Virus from Wild-Caught

Red-Capped Mangabeys (*Cercocebus torquatus*) from Cameroon: a New Primate T-Lymphotropic Virus Type 3 Subtype. *J. Virol.* **76**(1):259-268.

7. Sitbon, M., L. d'Auriol, H. Ellerbrok, C. Andre, J. Nishio, S. Perryman, F. Pozo, S. F. Hayes, K. Wehrly, P. Tambourin, F. Galibert, and B. Chesebro. 1991. Substitution of  
5 leucine for isoleucine in a sequence highly conserved among retroviral envelope surface glycoproteins attenuates the lytic effect of the Friend murine leukemia virus. *Proc Natl Acad Sci U S A.* **88**(13):5932-6.

## REVENDICATIONS

1. Utilisation de couples d'oligonucléotides dégénérés en orientation 5' et 3' issus des séquences nucléotidiques codant pour la région aminoterminal de la composante de surface  
5 (SU) des protéines d'enveloppe des virus des lymphomes / leucémies T chez les primates, regroupés sous la désignation PTLV, ces virus étant encore désignés HTLV chez l'homme et STLV chez le singe, à savoir de la région correspondant aux fragments protéiques délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des  
10 différentes souches des PTLV, ou de virus portant des séquences apparentées aux SU des PTLV,

pour la mise en œuvre de procédés de détection de toute souche de PTLV, à savoir de toute souche appartenant aux HTLV-1, HTLV-2, STLV-1, STLV-2, et STLV-3, ainsi que de toute souche de virus apparentés aux PTLV, à savoir de toute souche dont la séquence en  
15 acides aminés déduite de la séquence nucléotidique codant pour la région aminoterminal de la SU présente un taux d'homologie d'au moins environ 30% avec les séquences en acides aminés codées par les séquences nucléotidiques correspondantes chez les PTLV, notamment pour la détection de nouveaux variants des PTLV, ou de virus comportant des séquences apparentées aux SU des PTLV, lesdits procédés comprenant une étape d'amplification, à  
20 partir d'un échantillon biologique susceptible de contenir des PTLV, et à l'aide des oligonucléotides 5' et 3' dégénérés susmentionnés utilisés en tant qu'amorces, du nombre de copies de fragments nucléotidiques délimités en position 5' par l'oligonucléotide dégénéré en orientation 5', et en position 3' par l'oligonucléotide dégénéré en orientation 3', et une étape d'identification de la souche de PTLV contenue dans l'échantillon biologique à partir des  
25 fragments nucléotidiques amplifiés susmentionnés.

2. Utilisation de couples d'oligonucléotides dégénérés selon la revendication 1, caractérisée en ce que lesdits oligonucléotides sont choisis parmi ceux comprenant environ 15 à environ 30 nucléotides issus des séquences nucléotidiques codant pour des fragments  
30 protéiques délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43, ou la souche NRA de HTLV-2

représentée par SEQ ID NO : 45, ou la souche de STLV-3 représentée par SEQ ID NO : 47, lesdits oligonucléotides dégénérés comprenant un mélange d'oligonucléotides issus de séquences codant pour une région déterminée d'environ 5 à 10 acides aminés des protéines d'enveloppe des différents souches de PTLV, et qui diffèrent entre eux par substitution d'au moins un nucléotide par un autre de manière à ce que chaque oligonucléotide soit susceptible de coder pour la région déterminée susmentionnée issue des fragments protéiques des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de HTLV-1, ou la souche NRA de HTLV-2, ou la souche de STLV-3 susmentionnées.

10

3. Utilisation de couples d'oligonucléotides dégénérés selon la revendication 1 ou 2, comprenant environ 15 à environ 30 nucléotides issus de séquences nucléotidiques codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus de fragments protéiques délimités par les acides aminés situés aux positions 80 à 245, et plus particulièrement aux positions 83 à 241, de la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43.

15

4. Utilisation de couples d'oligonucléotides dégénérés selon l'une des revendications 1 à 3, issus de séquences nucléotidiques codant pour les fragments polypeptidiques 83-88, 140-145, 222-228, et 237-241, de la protéine d'enveloppe de la souche MT-2 de HTLV-1, à savoir les fragments suivants :

20

83-YL/VFPHW-88 SEQ ID NO : 1

140-NFTQ/REV-145 SEQ ID NO : 2

222-NYS/TCI/MVC-228 SEQ ID NO : 3

25

237-WHVLY-241 SEQ ID NO : 4

5. Utilisation selon l'une des revendications 1 à 4, d'oligonucléotides dégénérés en orientation 5' issus du brin ADN (+) codant pour :

- le fragment polypeptidique 83-88 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (I) suivante :

30

TAYBTNTTYCCNCAYTGG (I) SEQ ID NO : 5

dans laquelle :

Y représente C ou T,

B représente C, G ou T,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

PTLVE5'83a TAYBTNTTYCCNCACTGG SEQ ID NO : 6

5 PTLVE5'83b TAYBTNTTYCCNCATTGG SEQ ID NO : 7

Y, B et N étant tels que définis ci-dessus,

- le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (II) suivante :

AAATTYACNCARGAGT (II) SEQ ID NO : 8

10 dans laquelle :

Y représente C ou T,

R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

15 PTLVE5'140a AAATTYACNCAAGAAGT SEQ ID NO : 9

PTLVE5'140b AAATTYACNCAGGAAGT SEQ ID NO : 10

PTLVE5'140c AAATTYACNCAAGAGGT SEQ ID NO : 11

PTLVE5'140d AAATTYACNCAGGAGGT SEQ ID NO : 12

Y et N étant tels que définis ci-dessus.

20

6. Utilisation selon l'une des revendications 1 à 4, d'oligonucléotides dégénérés en orientation 3' issus du brin ADN (-) codant pour :

- le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (III) suivante :

25 NACYTCYTGNGTAAATT (III) SEQ ID NO : 13

dans laquelle :

Y représente C ou T,

R représente A ou G,

N représente A, C, G ou T,

30 telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'145a NACYTCYTGNGTAAAATT SEQ ID NO : 14

PTLVE3'145b NACYTCYTGNGTGAAATT SEQ ID NO : 15

PTLVE3'145c NACYTCYTGNGTAAAGTT SEQ ID NO : 16

PTLVE3'145d NACYTCYTGNGTGAAGTT

SEQ ID NO : 17

Y et N étant tels que définis ci-dessus,

- le fragment polypeptidique 222-228 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (IV) suivante :

5 RMNACNATRCANSWRTARTT (IV) SEQ ID NO : 18

dans laquelle :

R représente A ou G,

M représente A ou C,

S représente C ou G,

10 W représente A ou T,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'228a RMNACNATRCANSAATAATT SEQ ID NO : 19

PTLVE3'228b RMNACNATRCANSAGTAATT SEQ ID NO : 20

.5 PTLVE3'228c RMNACNATRCANSAATAGTT SEQ ID NO : 21

PTLVE3'228d RMNACNATRCANSAGTAGTT SEQ ID NO : 22

PTLVE3'228e RMNACNATRCANSTATAATT SEQ ID NO : 23

PTLVE3'228f RMNACNATRCANSTGTAATT SEQ ID NO : 24

PTLVE3'228g RMNACNATRCANSTATAGIT SEQ ID NO : 25

10 PTLVE3'228h RMNACNATRCANSTGTAGTT SEQ ID NO : 26

R, M, S, et N étant tels que définis ci-dessus,

- le fragment polypeptidique 237-241 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (V) suivante :

RTANARNACRTGCCA (V) SEQ ID NO : 27

5 dans laquelle :

R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'241a RTANARNACATGCCA SEQ ID NO : 28

0 PTLVE3'241b RTANARNACGTGCCA SEQ ID NO : 29

R, et N étant tels que définis ci-dessus.

7. Utilisation selon l'une des revendications 1 à 6, de couples d'oligonucléotides dégénérés choisis de telle manière que :

5 - les oligonucléotides dégénérés 5' correspondent à un mélange d'oligonucléotides 5' issus d'une même région nucléotidique déterminée comprenant environ 15 à environ 30 nucléotides issus du brin ADN (+) et codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus de fragments protéiques délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 135 à 150 des protéines d'enveloppe des différentes souches des PTLV, notamment codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus du fragment protéique délimité du côté N-terminal par l'acide aminé situé à la position 83 et du côté C-terminal par l'acide aminé situé à la position 145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides 5' étant tels qu'ils diffèrent entre eux par substitution d'au moins un nucléotide par un autre de manière à ce que chaque oligonucléotide soit susceptible de coder pour la région déterminée susmentionnée  
10 issue des fragments protéiques des protéines d'enveloppe des différentes souches des PTLV,

- les oligonucléotides dégénérés 3' correspondent à un mélange d'oligonucléotides 3' issus d'une même région nucléotidique déterminée comprenant environ 15 à environ 30 nucléotides issus du brin ADN (-) et codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus de fragments protéiques délimités du côté N-terminal par un acide aminé situé entre les positions 125 à 145, et du côté C-terminal par un acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des différentes souches des PTLV, notamment codant pour des fragments polypeptidiques d'environ 5 à environ 10 acides aminés issus du fragment protéique délimité du côté N-terminal par l'acide aminé situé à la position 140 et du côté C-terminal par l'acide aminé situé à la position 241 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides 3' étant tels qu'ils diffèrent entre eux par substitution d'au moins un nucléotide par un autre de manière à ce que chaque oligonucléotide soit susceptible de coder pour la région déterminée susmentionnée  
15 issue des fragments protéiques des protéines d'enveloppe des différentes souches des PTLV,

étant bien entendu que lesdites amorces 5' et 3' susmentionnées ne peuvent pas être  
20 complémentaires l'une de l'autre.

8. Utilisation de couples d'oligonucléotides dégénérés selon l'une des revendications 1 à 7, caractérisée en ce que les oligonucléotides dégénérés 5' sont choisis parmi les

oligonucléotides 5' de formules (I) et (II) selon la revendication 5, et en ce que les oligonucléotides dégénérés 3' sont choisis parmi les oligonucléotides 3' de formules (III) à (V) selon la revendication 6.

5 9. Utilisation de couples d'oligonucléotides dégénérés selon l'une des revendications 1 à 8, lesdits oligonucléotides étant choisis de manière à ce qu'ils permettent l'amplification, à partir d'un échantillon biologique susceptible de contenir de l'ADN de PTLV, de séquences nucléotidiques codant pour des fragments protéiques comprenant une séquence délimitée du côté N-terminal par l'acide aminé situé en position 89, et du côté C-terminal par l'acide aminé  
10 situé en position 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43, ou comprenant une séquence analogue comprise dans la protéine d'enveloppe d'une autre souche de PTLV que HTLV-1, telle que la séquence délimitée du côté N-terminal par l'acide aminé situé en position 85, et du côté C-terminal par l'acide aminé situé en position 135 de la protéine d'enveloppe de la souche NRA de HTLV-2 représentée  
15 par SEQ ID NO : 45, ou la séquence délimitée du côté N-terminal par l'acide aminé situé en position 88, et du côté C-terminal par l'acide aminé situé en position 144 de la protéine d'enveloppe de la souche de STL-3 représentée par SEQ ID NO : 47.

10 10. Utilisation de couples d'oligonucléotides dégénérés selon la revendication 9, caractérisée en ce que les oligonucléotides dégénérés 5' sont choisis parmi les oligonucléotides 5' de formule (I) selon la revendication 5, et en ce que les oligonucléotides dégénérés 3' sont choisis parmi les oligonucléotides 3' de formules (III) à (V) selon la revendication 6.

15 11. Utilisation de couples d'oligonucléotides dégénérés selon la revendication 9 ou 10, caractérisée en ce que :

- les oligonucléotides dégénérés 5' sont ceux de formule (I) suivante :

PTLVE5'83b TAYBTNTTYCCNCATTGG SEQ ID NO : 5

Y, B et N étant tels que définis dans la revendication 5,

- les oligonucléotides dégénérés 3' sont ceux de formule (III) suivante :

20 PTLVE3'145a NACYTCYTGNGTAAAATT SEQ ID NO : 14

Y et N étant tels que définis dans la revendication 6.

12. Oligonucléotides tels que définis dans l'une des revendications 1 à 6.



13. Oligonucléotides selon la revendication 12, correspondants :

- aux oligonucléotides dégénérés en orientation 5' issus du brin ADN (+) codant pour :

\* le fragment polypeptidique 83-88 de la protéine d'enveloppe de la souche MT-

2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (I) suivante :

TAYBTNTTYCCNCAYTGG (I) SEQ ID NO : 5

dans laquelle :

Y représente C ou T,

B représente C, G ou T,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

PTLVE5'83a TAYBTNTTYCCNCACTGG SEQ ID NO : 6

PTLVE5'83b TAYBTNTTYCCNCATTGG SEQ ID NO : 7

Y, B et N étant tels que définis ci-dessus,

\* le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche

MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (II)

suivante :

AAYTTYACNCARGARGT (II) SEQ ID NO : 8

dans laquelle :

Y représente C ou T,

R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 5' choisies parmi les suivantes :

PTLVE5'140a AAYTTYACNCAAGAAGT SEQ ID NO : 9

PTLVE5'140b AAYTTYACNCAGGAAGT SEQ ID NO : 10

PTLVE5'140c AAYTTYACNCAAGAGGT SEQ ID NO : 11

PTLVE5'140d AAYTTYACNCAGGAGGT SEQ ID NO : 12

Y et N étant tels que définis ci-dessus,

- aux oligonucléotides dégénérés en orientation 3' issus du brin ADN (-) codant pour :

\* le fragment polypeptidique 140-145 de la protéine d'enveloppe de la souche

MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (III)

suivante :

NACYTCYTGNGTRAARTT

(III)

SEQ ID NO : 13

dans laquelle :

Y représente C ou T,

R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'145a      NACYTCYTGNGTAAAATT      SEQ ID NO : 14

PTLVE3'145b      NACYTCYTGNGTGAAATT      SEQ ID NO : 15

PTLVE3'145c      NACYTCYTGNGTAAAGTT      SEQ ID NO : 16

PTLVE3'145d      NACYTCYTGNGTGAAGTT      SEQ ID NO : 17

Y et N étant tels que définis ci-dessus,

\* le fragment polypeptidique 222-228 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (IV) suivante :

RMNACNATRCANSWRTARTT

(IV) SEQ ID NO : 18

dans laquelle :

R représente A ou G,

M représente A ou C,

S représente C ou G,

W représente A ou T,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'228a      RMNACNATRCANSAATAATT      SEQ ID NO : 19

PTLVE3'228b      RMNACNATRCANSAGTAATT      SEQ ID NO : 20

PTLVE3'228c      RMNACNATRCANSAATAGTT      SEQ ID NO : 21

PTLVE3'228d      RMNACNATRCANSAGTAGTT      SEQ ID NO : 22

PTLVE3'228e      RMNACNATRCANSTATAATT      SEQ ID NO : 23

PTLVE3'228f      RMNACNATRCANSTGTAATT      SEQ ID NO : 24

PTLVE3'228g      RMNACNATRCANSTATAGTT      SEQ ID NO : 25

PTLVE3'228h      RMNACNATRCANSTGTAGTT      SEQ ID NO : 26

R, M, S, et N étant tels que définis ci-dessus,

\* le fragment polypeptidique 237-241 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, lesdits oligonucléotides étant choisis parmi ceux de formule (V) suivante :

RTANARNACRTGCCA (V) SEQ ID NO : 27

dans laquelle :

R représente A ou G,

N représente A, C, G ou T,

telles que les amorces oligonucléotidiques 3' choisies parmi les suivantes :

PTLVE3'241a RTANARNACATGCCA SEQ ID NO : 28

PTLVE3'241b RTANARNACGTGCCA SEQ ID NO : 29

R, et N étant tels que définis ci-dessus.

14. Procédé de détection de toute souche de PTLV, à savoir de toute souche appartenant aux HTLV-1, HTLV-2, STLV-1, STLV-2, et STLV-3, ainsi que de toute souche de virus comportant des séquences apparentées aux SU des PTLV, telles que définies dans la revendication 1, caractérisé en ce qu'il comprend :

- la mise en contact d'un couple d'oligonucléotides dégénérés 5' et 3' tels que définis dans l'une des revendications 1 à 13, avec l'ADN génomique ou l'ADN complémentaire dérivé à partir d'ARN extraits du contenu d'un échantillon biologique (tels que cellules sanguines, de moelle osseuse, biopsies, notamment de peau ou autres organes, ou frottis) susceptible de contenir des PTLV tels que définis ci-dessus,

- l'amplification de fragments d'ADN codant pour un fragment des protéines d'enveloppe des différentes souches des PTLV tel que défini dans la revendication 1 ou 9,

- la détection des fragments d'ADN amplifiés lors de l'étape précédente, cette détection pouvant être corrélée à la détection et le cas échéant à l'identification de PTLV tels que définis ci-dessus dans ledit échantillon biologique.

15. Procédé de détection de toute souche de PTLV selon la revendication 14, caractérisé en ce que l'étape d'amplification comprend la mise en œuvre de deux réactions d'amplification, la deuxième réaction étant effectuée sur un échantillon de produits obtenus dans le cadre de la première réaction à l'aide des mêmes oligonucléotides 5' que dans le cas de la première réaction, et d'oligonucléotides 3' différents de ceux utilisés dans la première réaction, à savoir des amorces 3' dites « nichées » hybridant avec une région située plus en

amont de la séquence codant pour la SU que les amorces 3' utilisées dans la première réaction.

16. Procédé de détection de toute souche de PTLV selon la revendication 14 ou 15, caractérisé en ce qu'il comprend :

- une première réaction d'amplification génique effectuée à l'aide de couples d'oligonucléotides dégénérés choisis parmi les couples :

\* oligonucléotides de formule (I)/oligonucléotides de formule (IV), ou

\* oligonucléotides de formule (I)/oligonucléotides de formule (V), ou

\* oligonucléotides de formule (II)/oligonucléotides de formule (V),

- et une deuxième étape d'amplification du nombre de copies de fragments d'ADN obtenus lors de l'étape précédente à l'aide de couples d'oligonucléotides dégénérés choisis respectivement parmi les couples :

\* oligonucléotides de formule (I)/oligonucléotides de formule (III), ou

\* oligonucléotides de formule (I)/oligonucléotides de formule (III ou IV), ou

\* oligonucléotides de formule (II)/oligonucléotides de formule (IV),

- la détection des fragments d'ADN amplifiés lors de l'étape précédente, cette détection pouvant être corrélée à la détection et le cas échéant à l'identification de PTLV dans l'échantillon biologique.

17. Procédé de détection de toute souche de PTLV selon l'une des revendications 14 à 16, caractérisé en ce qu'il comprend :

- une première réaction d'amplification génique effectuée à l'aide de couples d'oligonucléotides dégénérés choisis de telle manière que :

\* les oligonucléotides dégénérés 5' sont ceux de formule (I) suivante :

PTLVE5'83b TAYBTNTTYCCNCATTGG SEQ ID NO : 5

Y, B et N étant tels que définis dans la revendication 5,

\* les oligonucléotides dégénérés 3' sont ceux de formule (V) suivante :

PTLVE3'241b RTANARNACGTGCCA SEQ ID NO : 29

R, et N étant tels que définis dans la revendication 6,

- une deuxième réaction d'amplification génique effectuée à l'aide de couples d'oligonucléotides dégénérés choisis de telle manière que :

\* les oligonucléotides dégénérés 5' sont ceux de formule (I) suivante :

PTLVE5'83b TAYBTNTTYCCNCATTGG

SEQ ID NO : 5

Y, B et N étant tels que définis dans la revendication 5,

\* les oligonucléotides dégénérés 3' sont ceux de formule (III) suivante :

PTLVE3'145a NACYTCYTGNGTAAAATT

SEQ ID NO : 14

5 Y et N étant tels que définis dans la revendication 6.

18. Trousse ou kit pour la mise en œuvre d'un procédé de détection selon l'une des revendications 14 à 17, caractérisé en ce qu'il comprend un couple d'amorces tel que défini dans l'une des revendications 1 à 13, et, le cas échéant, les réactifs nécessaires à la mise en œuvre de la réaction d'amplification par PCR et pour la détection des fragments amplifiés.

19. Application du procédé de détection selon l'une des revendications 14 à 17 :

- au diagnostic de pathologies liées à une infection par un PTLV, ou par un virus comportant des séquences apparentées aux SU des PTLV, chez l'homme ou l'animal, telles que les hémopathies, les maladies auto-immunes, les maladies inflammatoires, les maladies dégénératives,

- au criblage et à l'identification de nouveaux agents infectieux chez l'homme ou l'animal, et plus particulièrement de nouvelles souches, ou variants, de virus susceptibles d'être classés dans les PTLV, ou de virus comportant des séquences apparentées aux SU des PTLV,

- au criblage de gènes de prédisposition ou de résistance à des pathologies chez l'homme ou l'animal liées à la présence des PTLV ou de séquences apparentées, ou à une infection par un PTLV, telles que les hémopathies, les maladies auto-immunes, les maladies dégénératives,

- au criblage ou à la conception de nouveaux agents thérapeutiques comprenant des séquences entières ou partielles des protéines d'enveloppe de nouveaux variants de PTLV ainsi détectés,

- au criblage ou à la conception de nouveaux vecteurs de thérapie cellulaire utilisant les propriétés de tropisme des séquences entières ou partielles des protéines d'enveloppe de nouveaux variants de PTLV ainsi détectés.

20. Variants de type HTLV-1 tels qu'obtenus par mise en œuvre d'un procédé de détection selon l'une des revendications 14 à 17, correspondants :

- au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 31 suivante :

I K K P N P N G G G Y Y L A S Y S D  
 P C S L K C P Y L G C Q S W T C P Y  
 5 T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu arginine (R) en position 94, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu proline (P) et un résidu leucine (L) indiqués en gras et soulignés,

et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 30 suivante :

ATT AAA AAG CCA AAC CCA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
 CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
 15 ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 281, C en position 302, et G en position 333, sont respectivement remplacés par C, T, et A indiqués en gras et soulignés.

- au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 33 suivante :

V K K P N R N G G G Y Y L A S Y S D  
 P C S L K C P Y L G C Q S W T C P Y  
 T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu isoleucine (I) en position 89, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu valine (V) et un résidu leucine (L) indiqués en gras et soulignés,

et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 32 suivante :

GTT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
 CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
 ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle A en position 266, C en position 302, et G en position 333, sont respectivement remplacés par G, T, et A indiqués en gras et soulignés.

- 5 - au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 35 suivante :

I	K	K	P	N	R	N	G	G	G	Y	Y	<u>L</u>	A	S	Y	S	D
P	C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y
T	G	A	V	S	S	P	Y	W	K	F	Q	Q	D	V			

- 10 à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu sérine (S) en position 101, est remplacé par un résidu leucine (L) indiqué en gras et souligné,

- et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle  
15 comprend la séquence SEQ ID NO : 34 suivante :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAA CAA GAT GTC

- à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés  
20 aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle C en position 302, G en position 333, et G en position 408, sont respectivement remplacés par T, A, et A indiqués en gras et soulignés.

- au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 37 suivante :

25	I	K	K	P	N	R	N	G	G	G	Y	Y	<u>L</u>	A	S	Y	S	D
	P	C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y
	T	G	<u>P</u>	V	S	S	P	Y	W	K	F	Q	Q	D	V			

- à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés  
30 aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu alanine (A) en position 127, est remplacé par un résidu proline (P) indiqué en gras et souligné,

- et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle  
comprend la séquence SEQ ID NO : 36 suivante :

35	GTT	AAA	AAG	CCA	AAC	CGA	AAT	GGC	GGA	GGC	TAT	TAT	<u>T</u> TA	GCC	TCT	TAT	TCA	GAC
	CCT	TGT	TCC	TTA	AAA	TGC	CCA	TAC	CTG	GGG	TGC	CAA	TCA	TGG	ACC	TGC	CCC	TAT

ACA GGA CCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 379, est remplacé par C indiqué en gras et souligné,

5 - au variant dont la protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID NO : 39 suivante :

I	K	K	P	N	R	N	G	G	G	Y	<u>H</u>	S	A	S	Y	S	D	P
C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y	<u>A</u>	G
A	V	S	S	P	Y	W	K	F	Q	Q	D	V	N	F	T	Q	E	V

10 à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu tyrosine (Y) en position 100, et le résidu thréonine (T) en position 125, sont respectivement remplacés par un résidu histidine (H) et un résidu alanine (A) indiqués en gras et soulignés,

15 et dont la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 38 suivante :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT CAT TCA GCC TCT TAT TCA  
 GAC CCT TGT TCC TTA AAG TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC  
 CCC TAT GCA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC  
 20 AAT TTT ACC CAG GAA GTA

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 435 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle T en position 298, A en position 373, T en position 426, A en position 429, et T en position 435, sont respectivement remplacés par C, G, C, G, et A  
 25 indiqués en gras et soulignés.

21. Variant de type HTLV-2 tel qu'obtenu par mise en œuvre d'un procédé de détection selon la revendication 14, caractérisé en ce que:

- sa protéine d'enveloppe est telle qu'elle comprend la séquence peptidique SEQ ID  
 30 NO : 41 suivante :

I	<u>R</u>	K	P	N	R	Q	G	L	G	Y	Y	S	P	S	Y	N	D
P	C	S	L	Q	C	P	Y	L	G	<u>S</u>	Q	S	W	T	C	P	Y
T	<u>A</u>	P	V	S	<u>T</u>	P	S	W	<u>N</u>	F	H	S	D	V			

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés  
 35 aux positions 85 à 135 de la protéine d'enveloppe de la souche prototype NRA de HTLV-2



(décrite par Lee et al., 1993. Virology 196, 57-69 ; n° d'accès Genbank L20734.1), dans laquelle les résidus lysine (K) en position 86, cystéine (C) en position 113, glycine (G) en position 122, sérine (S) en position 126, et lysine (K) en position 130, sont respectivement remplacés par les résidus arginine (R), sérine (S), alanine (A), thréonine (T), et asparagine (N) indiqués en gras et soulignés,

- la séquence nucléotidique codant pour sa protéine d'enveloppe est telle qu'elle comprend la séquence SEQ ID NO : 40 suivante :

ATA AGA AAG CCA AAC AGA CAG GGC CTA GGG TAC TAC TCG CCT TCC TAC AAT GAC  
CCT TGC TCG CTA CAA TGC CCC TAC TTG GGC TCC CAA TCA TGG ACA TGC CCA TAC  
10 ACG GCC CCC GTC TCC ACT CCA TCC TGG AAT TTT CAT TCA GAT GTA

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 253 à 405 de la séquence codant pour la protéine d'enveloppe de la souche NRA de HTLV-2, dans laquelle A en position 257, G en position 258, T en position 267, A en position 282, C en position 294, T en position 300, A en position 333, G en position 338, G en position 365, G en position 377, G en position 390, et C en position 396, sont respectivement remplacés par G, A, C, G, T, C, G, C, C, C, T, et T indiqués en gras et soulignés.

22. Polypeptides délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 230 à 245 des protéines d'enveloppe des différentes souches des PTLV, telles que la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43, ou de la souche NRA de HTLV-2 représentée par SEQ ID NO : 45, ou de la souche de STLV-3 représentée par SEQ ID NO : 47, ou de virus portant des séquences apparentées aux SU des PTLV, ou délimités du côté N-terminal par un acide aminé situé entre les positions 75 à 90, et du côté C-terminal par un acide aminé situé entre les positions 135 à 150 desdites protéines d'enveloppe des différentes souches des PTLV.

23. Polypeptides selon la revendication 22, choisis parmi :

- le polypeptide délimité du côté N-terminal par l'acide aminé situé à la position 83 ou 89, et du côté C-terminal par l'acide aminé situé à la position 139 ou 145, de la protéine d'enveloppe de la souche MT-2 de HTLV-1 représentée par SEQ ID NO : 43,

- le polypeptide délimité du côté N-terminal par l'acide aminé situé à la position 79 ou 85, et du côté C-terminal par l'acide aminé situé à la position 135 ou 141, de la protéine d'enveloppe de la souche NRA de HTLV-2 représentée par SEQ ID NO : 45,

- le polypeptide délimité du côté N-terminal par l'acide aminé situé à la position 82 ou 88, et du côté C-terminal par l'acide aminé situé à la position 138 ou 144, de la protéine d'enveloppe de la souche de STLV-3 représentée par SEQ ID NO : 47.

24. Polypeptides codés par les fragments d'ADN amplifiés dans le cadre du procédé de détection selon la revendication 14, des variants de type HTLV-1 à HTLV-2 selon l'une des revendications 17 à 21, caractérisés en ce qu'ils comprennent les séquences peptidiques suivantes :

- polypeptide 1 (SEQ ID NO : 31) :

10 I K K P N P N G G G Y Y L A S Y S D  
P C S L K C P Y L G C Q S W T C P Y  
T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu arginine (R) en position 94, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu proline (P) et un résidu leucine (L) indiqués en gras et soulignés,

-polypeptide 2 (SEQ ID NO : 33):

20 V K K P N R N G G G Y Y L A S Y S D  
P C S L K C P Y L G C Q S W T C P Y  
T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu isoleucine (I) en position 89, et le résidu sérine (S) en position 101, sont respectivement remplacés par un résidu valine (V) et un résidu leucine (L) indiqués en gras et soulignés,

- polypeptide 3 (SEQ ID NO : 35):

30 I K K P N R N G G G Y Y L A S Y S D  
P C S L K C P Y L G C Q S W T C P Y  
T G A V S S P Y W K F Q Q D V

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu sérine (S) en position 101, est remplacé par un résidu leucine (L) indiqué en gras et souligné.

- polypeptide 4 (SEQ ID NO : 37):

I	K	K	P	N	R	N	G	G	G	Y	Y	L	A	S	Y	S	D
P	C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y
T	G	<u>P</u>	V	S	S	P	Y	W	K	F	Q	Q	D	V			

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 139 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu alanine (A) en position 127, est remplacé par un résidu proline (P) indiqué en gras et souligné,

- polypeptide 5 (SEQ ID NO : 39) :

I	K	K	P	N	R	N	G	G	G	Y	<u>H</u>	S	A	S	Y	S	D
C	S	L	K	C	P	Y	L	G	C	Q	S	W	T	C	P	Y	<u>A</u>
A	V	S	S	P	Y	W	K	F	Q	Q	D	V	N	F	T	Q	E

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 89 à 145 de la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle le résidu tyrosine (Y) en position 100, et le résidu thréonine (T) en position 125, sont respectivement remplacés par un résidu histidine (H) et un résidu alanine (A) indiqués en gras et soulignés,

- polypeptide 6 (SEQ ID NO : 41) :

I	<u>R</u>	K	P	N	R	Q	G	L	G	Y	Y	S	P	S	Y	N	D
P	C	S	L	Q	C	P	Y	L	G	<u>S</u>	Q	S	W	T	C	P	Y
T	<u>A</u>	P	V	S	<u>T</u>	P	S	W	<u>N</u>	F	H	S	D	V			

à savoir une séquence correspondant à la séquence délimitée par les acides aminés situés aux positions 85 à 135 de la protéine d'enveloppe de la souche prototype NRA de HTLV-2, dans laquelle les résidus lysine (K) en position 86, cystéine (C) en position 113, glycine (G) en position 122, sérine (S) en position 126, et lysine (K) en position 130, sont respectivement remplacés par les résidus arginine (R), sérine (S), alanine (A), thréonine (T), et asparagine (N) indiqués en gras et soulignés.

25. Acides nucléiques caractérisés en ce qu'ils codent pour un polypeptide selon l'une des revendications 22 à 24.

26. Acides nucléiques selon la revendication 25, comprenant les séquences nucléotidiques suivantes :

- acide nucléique 1 a (SEQ ID NO : 30):

ATT	AAA	AAG	CCA	AAC	<u>CCA</u>	AAT	GGC	GGA	GGC	TAT	TAT	<u>TTA</u>	GCC	TCT	TAT	TCA	GAC
CCT	TGT	TCC	TTA	<u>AAA</u>	TGC	CCA	TAC	CTG	GGG	TGC	CAA	TCA	TGG	ACC	TGC	CCC	TAT

ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 281, C en position 302, et G en position 333, sont respectivement remplacés par C, T, et A indiqués en gras et soulignés,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 1 selon la revendication 24,

- acide nucléique 2 a (SEQ ID NO : 32):

GTT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle A en position 266, C en position 302, et G en position 333, sont respectivement remplacés par G, T, et A indiqués en gras et soulignés,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 2 la revendication 24,

- acide nucléique 3 a (SEQ ID NO : 34) :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAA CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle C en position 302, G en position 333, et G en position 408, sont respectivement remplacés par T, A, et A indiqués en gras et soulignés,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 3 la revendication 24,

- acide nucléique 4 a (SEQ ID NO : 36):

GTT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TTA GCC TCT TAT TCA GAC  
CCT TGT TCC TTA AAA TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC CCC TAT  
ACA GGA CCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 417 de la séquence codant pour la protéine d'enveloppe de la souche MT-2 de HTLV-1, dans laquelle G en position 379, est remplacé par C indiqué en gras et souligné,

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 4 selon la revendication 24,

- acide nucléique 5 a (SEQ ID NO : 38) :

ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT CAT TCA GCC TCT TAT TCA  
 5 GAC CCT TGT TCC TTA AAG TGC CCA TAC CTG GGG TGC CAA TCA TGG ACC TGC  
 CCC TAT GCA GGA GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG CAA GAT GTC  
 AAT TTT ACC CAG GAA GTA

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 265 à 435 de la séquence codant pour la protéine d'enveloppe de la souche MT-  
 10 2 de HTLV-1, dans laquelle T en position 298, A en position 373, T en position 426, A en position 429, et T en position 435, sont respectivement remplacés par C, G, C, G, et A indiqués en gras et soulignés.

ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 5 selon la revendication 24,

15 - acide nucléique 6 a (SEQ ID NO : 40):

ATA AGA AAG CCA AAC AGA CAG GGC CTA GGG TAC TAC TCG CCT TCC TAC AAT GAC  
 CCT TGC TCG CTA CAA TGC CCC TAC TTG GGC TCC CAA TCA TGG ACA TGC CCA TAC  
 ACG GCC CCC GTC TCC ACT CCA TCC TGG AAT TTT CAT TCA GAT GTA

à savoir une séquence correspondant à la séquence délimitée par les nucléotides situés aux positions 253 à 405 de la séquence codant pour la protéine d'enveloppe de la souche NRA  
 20 de HTLV-2, dans laquelle A en position 257, G en position 258, T en position 267, A en position 282, C en position 294, T en position 300, A en position 333, G en position 338, G en position 365, G en position 377, G en position 390, et C en position 396, sont respectivement remplacés par G, A, C, G, T, C, G, C, C, C, T, et T indiqués en gras et soulignés,

25 ou toute séquence nucléotidique dérivée par dégénérescence du code génétique et codant pour le polypeptide 6 selon la revendication 24.

27. Anticorps polyclonaux ou monoclonaux dirigés contre un variant selon la revendication 20 ou 21, ou contre un polypeptide selon l'une des revendications 22 à 24, tels  
 30 qu'obtenus par immunisation d'un animal approprié avec un polypeptide susmentionné.

28. Composition pharmaceutique, notamment vaccins ou vecteurs thérapeutiques, comprenant, un polypeptide selon l'une des revendications 22 à 24, ou un acide nucléique

selon la revendication 25 ou 26, ou des anticorps selon la revendication 27, le cas échéant en association avec un véhicule pharmaceutiquement acceptable.

## LISTE DE SEQUENCES

<110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

<120> OLIGONUCLEOTIDES ISSUS DES SEQUENCES CODANT POUR  
LA COMPOSANTE DE SURFACE DES PROTEINES D'ENVELOPPE  
DES PTLV ET LEURS UTILISATIONS

<130> IFB 02 BC CNR PTLV

<140> FR 02/05001

<141> 2002-04-22

<160> 47

<170> PatentIn version 3.1

<210> 1

<211> 6

<212> PRT

<213> Human T-cell lymphotropic virus type 1

<220>

<221> misc\_feature

<222> (2)..(2)

<223> L ou V

<400> 1  
Tyr Xaa Phe Pro His Trp  
1 5

<210> 2

<211> 6

<212> PRT

<213> Human T-cell lymphotropic virus type 1

<220>

<221> misc\_feature

<222> (4)..(4)

<223> Q ou R

<400> 2  
Asn Phe Thr Xaa Glu Val  
1 5

<210> 3

<211> 7

<212> PRT

<213> Human T-cell lymphotropic virus type 1

<220>

<221> misc\_feature

<222> (3)..(3)

<223> S ou T

<220>

<221> misc\_feature

<222> (5)..(5)

<223> I ou M

<400> 3

Asn Tyr Xaa Cys Xaa Val Cys

1 5

<210> 4

<211> 5

<212> PRT

<213> Human T-cell lymphotropic virus type 1

<400> 4

Trp His Val Leu Tyr

1 5

<210> 5

<211> 18

<212> ADN

<213> séquence artificielle

<220>

<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 83-88 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>

<221> misc\_feature

<222> (3)..(3)

<223> C ou T

<220>

<221> misc\_feature

<222> (4)..(4)

<223> C, G ou T

<220>

<221> misc\_feature

<222> (6)..(6)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (9)..(9)

<223> C ou T

<220>

<221> misc\_feature

<222> (12)..(12)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (15)..(15)

<223> C ou T

<400> 5

tanntnttnc cncantgg

18

<210> 6



<211> 18  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 83-88 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<400> 6  
tanntnttnc cncactgg

18

<210> 7  
<211> 18  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 83-88 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (9)..(9)

<223> C ou T

<220>

<221> misc\_feature

<222> (12)..(12)

<223> A, C, G ou T

<400> 7

tanntnttnc cncattgg

18

<210> 8

<211> 17

<212> ADN

<213> séquence artificielle

<220>

<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>

<221> misc\_feature

<222> (3)..(3)

<223> C ou T

<220>

<221> misc\_feature

<222> (6)..(6)

<223> C ou T

<220>

<221> misc\_feature

<222> (9)..(9)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (12)..(12)

<223> A ou G

<220>

<221> misc\_feature

<222> (15)..(15)

<223> A ou G

<400> 8

aannttnacnc angangt

17

<210> 9

<211> 17

<212> ADN

<213> séquence artificielle

<220>

5/32

<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (3)..(3)

&lt;223&gt; C ou T

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (6)..(6)

&lt;223&gt; C ou T

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (9)..(9)

&lt;223&gt; A, C, G ou T

&lt;400&gt; 9

aanttnacnc aagaagt

17

&lt;210&gt; 10

&lt;211&gt; 17

&lt;212&gt; ADN

&lt;213&gt; séquence artificielle

&lt;220&gt;

<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (3)..(3)

&lt;223&gt; C ou T

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (6)..(6)

&lt;223&gt; C ou T

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (9)..(9)

&lt;223&gt; A, C, G ou T

&lt;400&gt; 10

aanttnacnc aggaagt

17

&lt;210&gt; 11

&lt;211&gt; 17

&lt;212&gt; ADN

&lt;213&gt; séquence artificielle

&lt;220&gt;

<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A, C, G ou T

<400> 11  
aanttnacnc aagaggt

17

<210> 12  
<211> 17  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 5' issu du brin ADN (+) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A, C, G ou T

<400> 12  
aanttnacnc aggaggt

17

<210> 13  
<211> 18  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature

<222> (1)..(1)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (10)..(10)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (16)..(16)  
<223> A ou G

<400> 13  
nacntcntgn gtnaantt

18

<210> 14  
<211> 18  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (10)..(10)  
<223> A, C, G ou T

<400> 14  
nacntcntgn gtaaaatt

18

<210> 15  
<211> 18  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> C ou T

<220>  
<221> misc\_feature  
<222> (10)..(10)  
<223> A, C, G ou T

<400> 15  
nacntcntgn gtgaaatt

18

<210> 16  
<211> 18  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> C ou T

<220>  
<221> misc\_feature

<222> (7)..(7)

<223> C ou T

<220>

<221> misc\_feature

<222> (10)..(10)

<223> A, C, G ou T

<400> 16

nacntcntgn gtaaagtt

18

<210> 17

<211> 18

<212> ADN

<213> séquence artificielle

<220>

<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 140-145 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>

<221> misc\_feature

<222> (1)..(1)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (4)..(4)

<223> C ou T

<220>

<221> misc\_feature

<222> (7)..(7)

<223> C ou T

<220>

<221> misc\_feature

<222> (10)..(10)

<223> A, C, G ou T

<400> 17

nacntcntgn gtgaagtt

18

<210> 18

<211> 20

<212> ADN

<213> séquence artificielle

<220>

<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>

<221> misc\_feature

<222> (1)..(1)

<223> A ou G

<220>  
<221> misc\_feature  
<222> (2)..(2)  
<223> A ou C

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> C ou G

<220>  
<221> misc\_feature  
<222> (14)..(14)  
<223> A ou T

<220>  
<221> misc\_feature  
<222> (15)..(15)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (18)..(18)  
<223> A ou G

<400> 18  
nnnacnatnc annwntanttt

20

<210> 19  
<211> 20  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature



<222> (1)..(1)

<223> A ou G

<220>

<221> misc\_feature

<222> (2)..(2)

<223> A ou C

<220>

<221> misc\_feature

<222> (3)..(3)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (6)..(6)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (9)..(9)

<223> A ou G

<220>

<221> misc\_feature

<222> (12)..(12)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (13)..(13)

<223> C ou G

<400> 19

nnnacnatnc annaataatt

20

<210> 20

<211> 20

<212> ADN

<213> séquence artificielle

<220>

<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>

<221> misc\_feature

<222> (1)..(1)

<223> A ou G

<220>

<221> misc\_feature

<222> (2)..(2)

<223> A ou C

<220>

<221> misc\_feature

<222> (3)..(3)

<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> C ou G

<400> 20  
nnnacnatnc annagtaatt

20

<210> 21  
<211> 20  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (2)..(2)  
<223> A ou C

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A ou G

<220>

<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> C ou G

<400> 21  
nnnacnatnc annaatagtt

20

<210> 22  
<211> 20  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (2)..(2)  
<223> A ou C

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> C ou G

<400> 22  
nnnacnatnc annagtagtt

20

<210> 23  
<211> 20  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (2)..(2)  
<223> A ou C

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> C ou G

<400> 23  
nnnacnatnc anntataatt

20

<210> 24  
<211> 20  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>

<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (2)..(2)  
<223> A ou C

<220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> C ou G

<400> 24  
nnnacnatnc anntgtaatt

20

<210> 25  
<211> 20  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (2)..(2)  
<223> A ou C

<220>  
<221> misc\_feature  
<222> (3)..(3)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (6)..(6)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (9)..(9)

<223> A ou G

<220>

<221> misc\_feature

<222> (12)..(12)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (13)..(13)

<223> C ou G

<400> 25

nnnacnatnc anntatagtt

20

<210> 26

<211> 20

<212> ADN

<213> séquence artificielle

<220>

<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 222-228 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>

<221> misc\_feature

<222> (1)..(1)

<223> A ou G

<220>

<221> misc\_feature

<222> (2)..(2)

<223> A ou C

<220>

<221> misc\_feature

<222> (3)..(3)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (6)..(6)

<223> A, C, G ou T

<220>

<221> misc\_feature

<222> (9)..(9)

<223> A ou G

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> C ou G

<400> 26  
nnnacnatnc anntgtagtt

20

<210> 27  
<211> 15  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 237-241 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (10)..(10)  
<223> A ou G

<400> 27  
ntanannacn tgcca

15

<210> 28  
<211> 15  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 237-241 de la protéine

d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> A, C, G ou T

<400> 28  
ntanannaca tgcca

15

<210> 29  
<211> 15  
<212> ADN  
<213> séquence artificielle

<220>  
<223> oligonucléotide 3' issu du brin ADN (-) correspondant  
au fragment polypeptidique 237-241 de la protéine  
d'enveloppe de la souche MT-2 de HTLV-1

<220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> A, C, G ou T

<220>  
<221> misc\_feature  
<222> (6)..(6)  
<223> A ou G

<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> A, C, G ou T

<400> 29  
ntanannacg tgcca

15

<210> 30



<211> 153  
 <212> ADN  
 <213> Human T-cell lymphotropic virus type 1

<220>  
 <221> CDS  
 <222> (1)..(153)  
 <223>

<400> 30  
 att aaa aag cca aac cca aat ggc gga ggc tat tat tta gcc tct tat 48  
 Ile Lys Lys Pro Asn Pro Asn Gly Gly Gly Tyr Tyr Leu Ala Ser Tyr  
 1 5 10 15  
 tca gac cct tgt tcc tta aaa tgc cca tac ctg ggg tgc caa tca tgg 96  
 Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
 20 25 30  
 acc tgc ccc tat aca gga gcc gtc tcc agc ccc tac tgg aag ttt cag 144  
 Thr Cys Pro Tyr Thr Gly Ala Val Ser Ser Pro Tyr Trp Lys Phe Gln  
 35 40 45  
 caa gat gtc 153  
 Gln Asp Val  
 50

<210> 31  
 <211> 51  
 <212> PRT  
 <213> Human T-cell lymphotropic virus type 1

<400> 31  
 ile lys lys pro asn pro asn gly gly gly tyr tyr leu ala ser tyr  
 1 5 10 15  
 ser asp pro cys ser leu lys cys pro tyr leu gly cys gln ser trp  
 20 25 30  
 thr cys pro tyr thr gly ala val ser ser pro tyr trp lys phe gln  
 35 40 45  
 gln asp val  
 50

<210> 32  
 <211> 153  
 <212> ADN  
 <213> Human T-cell lymphotropic virus type 1

<220>  
 <221> CDS  
 <222> (1)..(153)  
 <223>

<400> 32  
 gtt aaa aag cca aac cga aat ggc gga ggc tat tat tta gcc tct tat 48  
 Val Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr Tyr Leu Ala Ser Tyr  
 1 5 10 15

20/32

tca gac cct tgt tcc tta aaa tgc cca tac ctg ggg tgc caa tca tgg 96  
 Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
                   20                                  25                                  30

acc tgc ccc tat aca gga gcc gtc tcc agc ccc tac tgg aag ttt cag 144  
 Thr Cys Pro Tyr Thr Gly Ala Val Ser Ser Pro Tyr Trp Lys Phe Gln  
                   35                                  40                                  45

caa gat gtc 153  
 Gln Asp Val  
           50

&lt;210&gt; 33

&lt;211&gt; 51

&lt;212&gt; PRT

&lt;213&gt; Human T-cell lymphotropic virus type 1

&lt;400&gt; 33

Val Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr Tyr Leu Ala Ser Tyr  
 1                                  5                                  10                                  15

Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
                   20                                  25                                  30

Thr Cys Pro Tyr Thr Gly Ala Val Ser Ser Pro Tyr Trp Lys Phe Gln  
                   35                                  40                                  45

Gln Asp Val  
           50

&lt;210&gt; 34

&lt;211&gt; 153

&lt;212&gt; ADN

&lt;213&gt; Human T-cell lymphotropic virus type 1

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (153)

&lt;223&gt;

&lt;400&gt; 34

att aaa aag cca aac cga aat ggc gga ggc tat tat tta gcc tct tat 48  
 Ile Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr Tyr Leu Ala Ser Tyr  
 1                                  5                                  10                                  15

tca gac cct tgt tcc tta aaa tgc cca tac ctg ggg tgc caa tca tgg 96  
 Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
                   20                                  25                                  30

acc tgc ccc tat aca gga gcc gtc tcc agc ccc tac tgg aag ttt caa 144  
 Thr Cys Pro Tyr Thr Gly Ala Val Ser Ser Pro Tyr Trp Lys Phe Gln  
                   35                                  40                                  45

caa gat gtc 153  
 Gln Asp Val  
           50

&lt;210&gt; 35

<211> 51  
 <212> PRT  
 <213> Human T-cell lymphotropic virus type 1

<400> 35  
 Ile Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr Tyr Leu Ala Ser Tyr  
 1 5 10 15  
 Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
 20 25 30  
 Thr Cys Pro Tyr Thr Gly Ala Val Ser Ser Pro Tyr Trp Lys Phe Gln  
 35 40 45  
 Gln Asp Val  
 50

<210> 36  
 <211> 153  
 <212> ADN  
 <213> Human T-cell lymphotropic virus type 1

<220>  
 <221> CDS  
 <222> (1)..(153)  
 <223>

<400> 36  
 gtt aaa aag cca aac cga aat ggc gga ggc tat tat tta gcc tct tat 48  
 Val Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr Tyr Leu Ala Ser Tyr  
 1 5 10 15  
 tca gac cct tgt tcc tta aaa tgc cca tac ctg ggg tgc caa tca tgg 96  
 Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
 20 25 30  
 acc tgc ccc tat aca gga ccc gtc tcc agc ccc tac tgg aag ttt cag 144  
 Thr Cys Pro Tyr Thr Gly Pro Val Ser Ser Pro Tyr Trp Lys Phe Gln  
 35 40 45  
 caa gat gtc 153  
 Gln Asp Val  
 50

<210> 37  
 <211> 51  
 <212> PRT  
 <213> Human T-cell lymphotropic virus type 1

<400> 37  
 Val Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr Tyr Leu Ala Ser Tyr  
 1 5 10 15  
 Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
 20 25 30  
 Thr Cys Pro Tyr Thr Gly Pro Val Ser Ser Pro Tyr Trp Lys Phe Gln  
 35 40 45

Gln Asp Val  
50

<210> 38  
<211> 171  
<212> ADN  
<213> Human T-cell lymphotropic virus type 1

<220>  
<221> CDS  
<222> (1)..(171)  
<223>

<400> 38  
att aaa aag cca aac cga aat ggc gga ggc tat cat tca gcc tct tat 48  
Ile Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr His Ser Ala Ser Tyr  
1 5 10 15  
tca gac cct tgt tcc tta aag tgc cca tac ctg ggg tgc caa tca tgg 96  
Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
20 25 30  
acc tgc ccc tat gca gga gcc gtc tcc agc ccc tac tgg aag ttt cag 144  
Thr Cys Pro Tyr Ala Gly Ala Val Ser Ser Pro Tyr Trp Lys Phe Gln  
35 40 45  
caa gat gtc aat ttt acc cag gaa gta 171  
Gln Asp Val Asn Phe Thr Gln Glu Val  
50 55

<210> 39  
<211> 57  
<212> PRT  
<213> Human T-cell lymphotropic virus type 1

<400> 39  
Ile Lys Lys Pro Asn Arg Asn Gly Gly Gly Tyr His Ser Ala Ser Tyr  
1 5 10 15  
Ser Asp Pro Cys Ser Leu Lys Cys Pro Tyr Leu Gly Cys Gln Ser Trp  
20 25 30  
Thr Cys Pro Tyr Ala Gly Ala Val Ser Ser Pro Tyr Trp Lys Phe Gln  
35 40 45  
Gln Asp Val Asn Phe Thr Gln Glu Val  
50 55

<210> 40  
<211> 153  
<212> ADN  
<213> Human T-cell lymphotropic virus type 2

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (153)

&lt;223&gt;

&lt;400&gt; 40

ata	aga	aag	cca	aac	aga	cag	ggc	cta	ggg	tac	tac	tcg	cct	tcc	tac	48
Ile	Arg	Lys	Pro	Asn	Arg	Gln	Gly	Leu	Gly	Tyr	Tyr	Ser	Pro	Ser	Tyr	
1				5					10					15		

aat	gac	cct	tgc	tcg	cta	caa	tgc	ccc	tac	ttg	ggc	tcc	caa	tca	tgg	96
Asn	Asp	Pro	Cys	Ser	Leu	Gln	Cys	Pro	Tyr	Leu	Gly	Ser	Gln	Ser	Trp	
			20					25					30			

aca	tgc	cca	tac	acg	gcc	ccc	gtc	tcc	act	cca	tcc	tgg	aat	ttt	cat	144
Thr	Cys	Pro	Tyr	Thr	Ala	Pro	Val	Ser	Thr	Pro	Ser	Trp	Asn	Phe	His	
			35				40					45				

tca	gat	gta														153
Ser	Asp	Val														
			50													

&lt;210&gt; 41

&lt;211&gt; 51

&lt;212&gt; PRT

&lt;213&gt; Human T-cell lymphotropic virus type 2

&lt;400&gt; 41

Ile	Arg	Lys	Pro	Asn	Arg	Gln	Gly	Leu	Gly	Tyr	Tyr	Ser	Pro	Ser	Tyr
1				5					10					15	

Asn	Asp	Pro	Cys	Ser	Leu	Gln	Cys	Pro	Tyr	Leu	Gly	Ser	Gln	Ser	Trp
			20					25					30		

Thr	Cys	Pro	Tyr	Thr	Ala	Pro	Val	Ser	Thr	Pro	Ser	Trp	Asn	Phe	His
			35				40					45			

Ser	Asp	Val													
			50												

&lt;210&gt; 42

&lt;211&gt; 924

&lt;212&gt; ADN

&lt;213&gt; Human T-cell lymphotropic virus type 1

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (924)

&lt;223&gt;

&lt;400&gt; 42

atg	ggg	aag	ttt	ctc	gcc	act	ttg	att	tta	ttc	ttc	cag	ttc	tgc	ccc	48
Met	Gly	Lys	Phe	Leu	Ala	Thr	Leu	Ile	Leu	Phe	Phe	Gln	Phe	Cys	Pro	
1				5					10					15		

ctc	atc	ctc	ggg	gat	tac	agc	ccc	agc	tgc	tgt	act	ctc	aca	att	gga	96
Leu	Ile	Leu	Gly	Asp	Tyr	Ser	Pro	Ser	Cys	Cys	Thr	Leu	Thr	Ile	Gly	
			20					25					30			

24/32

gtc tcc tca tac cac tct aaa ccc tgc aat cct gcc cag cca gtt tgt Val Ser Ser Tyr His Ser Lys Pro Cys Asn Pro Ala Gln Pro Val Cys 35 40 45	144
tgc tgg acc ctc gac ctg ctg gcc ctt tca gcg gat cag gcc cta cag Ser Trp Thr Leu Asp Leu Leu Ala Leu Ser Ala Asp Gln Ala Leu Gln 50 55 60	192
ccc ccc tgc cct aat cta gta agt tac tcc agc tac cat gcc acc tat Pro Pro Cys Pro Asn Leu Val Ser Tyr Ser Ser Tyr His Ala Thr Tyr 65 70 75 80	240
tcc cta tat cta ttc cct cat tgg att aaa aag cca aac cga aat ggc Ser Leu Tyr Leu Phe Pro His Trp Ile Lys Lys Pro Asn Arg Asn Gly 85 90 95	288
gga ggc tat tat tca gcc tct tat tca gac cct tgt tcc tta aag tgc Gly Gly Tyr Tyr Ser Ala Ser Tyr Ser Asp Pro Cys Ser Leu Lys Cys 100 105 110	336
cca tac ctg ggg tgc caa tca tgg acc tgc ccc tat aca gga gcc gtc Pro Tyr Leu Gly Cys Gln Ser Trp Thr Cys Pro Tyr Thr Gly Ala Val 115 120 125	384
tcc agc ccc tac tgg aag ttt cag caa gat gtc aat ttt act caa gaa Ser Ser Pro Tyr Trp Lys Phe Gln Gln Asp Val Asn Phe Thr Gln Glu 130 135 140	432
gtt tca cgc ctc aat att aat ctc cat ttt tca aaa tgc ggt ttt ccc Val Ser Arg Leu Asn Ile Asn Leu His Phe Ser Lys Cys Gly Phe Pro 145 150 155 160	480
ttc tcc ctt cta gtc gac gct cca gga tat gac ccc atc tgg ttc ctt Phe Ser Leu Leu Val Asp Ala Pro Gly Tyr Asp Pro Ile Trp Phe Leu 165 170 175	528
aat acc gaa ccc agc caa ctg cct ccc acc gcc cct cct cta ctc ccc Asn Thr Glu Pro Ser Gln Leu Pro Pro Thr Ala Pro Pro Leu Leu Pro 180 185 190	576
cac tct aac cta gac cac atc ctc gag ccc tct ata cca tgg aaa tca His Ser Asn Leu Asp His Ile Leu Glu Pro Ser Ile Pro Trp Lys Ser 195 200 205	624
aaa ctc ctg acc ctt gtc cag tta acc cta caa agc act aat tat act Lys Leu Leu Thr Leu Val Gln Leu Thr Leu Gln Ser Thr Asn Tyr Thr 210 215 220	672
tgc att gtc tgt atc gat cgt gcc agc cta tcc act tgg cac gtc cta Cys Ile Val Cys Ile Asp Arg Ala Ser Leu Ser Thr Trp His Val Leu 225 230 235 240	720
tac tct ccc aac gtc tct gtt cca tcc tct tct tct acc ccc ctc ctt Tyr Ser Pro Asn Val Ser Val Pro Ser Ser Ser Ser Thr Pro Leu Leu 245 250 255	768
tac cca tgc tta gcg ctt cca gcc ccc cac ctg acg tta cca ttt aac Tyr Pro Ser Leu Ala Leu Pro Ala Pro His Leu Thr Leu Pro Phe Asn 260 265 270	816
tgg acc cac tgc ttt gac ccc cag att caa gct ata gtc tcc tcc ccc	864

25/32

Trp Thr His Cys Phe Asp Pro Gln Ile Gln Ala Ile Val Ser Ser Pro  
 275 280 285

tgt cat aac tcc ctc atc ctg ccc ccc ttt tcc ttg tca cct gtt ccc 912  
 Cys His Asn Ser Leu Ile Leu Pro Pro Phe Ser Leu Ser Pro Val Pro  
 290 295 300

acc cta gga tcc 924  
 Thr Leu Gly Ser  
 305

<210> 43  
 <211> 308  
 <212> PRT  
 <213> Human T-cell lymphotropic virus type 1

<400> 43

Met Gly Lys Phe Leu Ala Thr Leu Ile Leu Phe Phe Gln Phe Cys Pro  
 1 5 10 15

Leu Ile Leu Gly Asp Tyr Ser Pro Ser Cys Cys Thr Leu Thr Ile Gly  
 20 25 30

Val Ser Ser Tyr His Ser Lys Pro Cys Asn Pro Ala Gln Pro Val Cys  
 35 40 45

Ser Trp Thr Leu Asp Leu Leu Ala Leu Ser Ala Asp Gln Ala Leu Gln  
 50 55 60

Pro Pro Cys Pro Asn Leu Val Ser Tyr Ser Ser Tyr His Ala Thr Tyr  
 65 70 75 80

Ser Leu Tyr Leu Phe Pro His Trp Ile Lys Lys Pro Asn Arg Asn Gly  
 85 90 95

Gly Gly Tyr Tyr Ser Ala Ser Tyr Ser Asp Pro Cys Ser Leu Lys Cys  
 100 105 110

Pro Tyr Leu Gly Cys Gln Ser Trp Thr Cys Pro Tyr Thr Gly Ala Val  
 115 120 125

Ser Ser Pro Tyr Trp Lys Phe Gln Gln Asp Val Asn Phe Thr Gln Glu  
 130 135 140

Val Ser Arg Leu Asn Ile Asn Leu His Phe Ser Lys Cys Gly Phe Pro  
 145 150 155 160

Phe Ser Leu Leu Val Asp Ala Pro Gly Tyr Asp Pro Ile Trp Phe Leu  
 165 170 175

Asn Thr Glu Pro Ser Gln Leu Pro Pro Thr Ala Pro Pro Leu Leu Pro  
 180 185 190

His Ser Asn Leu Asp His Ile Leu Glu Pro Ser Ile Pro Trp Lys Ser  
 195 200 205

Lys Leu Leu Thr Leu Val Gln Leu Thr Leu Gln Ser Thr Asn Tyr Thr  
 210 215 220

Cys Ile Val Cys Ile Asp Arg Ala Ser Leu Ser Thr Trp His Val Leu  
 225 230 235 240

Tyr Ser Pro Asn Val Ser Val Pro Ser Ser Ser Ser Thr Pro Leu Leu  
 245 250 255

Tyr Pro Ser Leu Ala Leu Pro Ala Pro His Leu Thr Leu Pro Phe Asn  
 260 265 270

Trp Thr His Cys Phe Asp Pro Gln Ile Gln Ala Ile Val Ser Ser Pro  
 275 280 285

Cys His Asn Ser Leu Ile Leu Pro Pro Phe Ser Leu Ser Pro Val Pro  
 290 295 300

Thr Leu Gly Ser  
 305

<210> 44  
 <211> 912  
 <212> ADN  
 <213> Human T-cell lymphotropic virus type 2

<220>  
 <221> CDS  
 <222> (1)..(912)  
 <223>

<400> 44  
 atg ggt aac gtt ttc ttc cta ctt tta ttc agt ctc aca cac ttc cca 48  
 Met Gly Asn Val Phe Phe Leu Leu Leu Phe Ser Leu Thr His Phe Pro  
 1 5 10 15

cca gtc cag cag agc cga tgc aca ctc acg gtt ggt att tcc tcc tac 96  
 Pro Val Gln Gln Ser Arg Cys Thr Leu Thr Val Gly Ile Ser Ser Tyr  
 20 25 30

cac tcc agc ccc tgt agc cca acc caa ccc gtc tgc acg tgg aac ctc 144  
 His Ser Ser Pro Cys Ser Pro Thr Gln Pro Val Cys Thr Trp Asn Leu  
 35 40 45



gac	ctt	aat	tcc	cta	acg	acg	gac	cag	cga	cta	cat	ccc	ccc	tgc	cct	192
Asp	Leu	Asn	Ser	Leu	Thr	Thr	Asp	Gln	Arg	Leu	His	Pro	Pro	Cys	Pro	
50						55				27/32	60					
aac	cta	att	act	tac	tct	ggc	ttc	cac	aaa	act	tat	tcc	tta	tac	tta	240
Asn	Leu	Ile	Thr	Tyr	Ser	Gly	Phe	His	Lys	Thr	Tyr	Ser	Leu	Tyr	Leu	
65					70					75					80	
ttc	cca	cat	tgg	ata	aag	aag	cca	aat	aga	cag	ggc	cta	gga	tac	tac	288
Phe	Pro	His	Trp	Ile	Lys	Lys	Pro	Asn	Arg	Gln	Gly	Leu	Gly	Tyr	Tyr	
				85					90					95		
tcg	ccc	tcc	tat	aat	gac	cct	tgc	tcg	cta	caa	tgc	ccc	tac	tta	ggc	336
Ser	Pro	Ser	Tyr	Asn	Asp	Pro	Cys	Ser	Leu	Gln	Cys	Pro	Tyr	Leu	Gly	
			100					105					110			
tgc	caa	tca	tgg	aca	tgc	cca	tac	acg	ggc	ccc	gtc	tcc	agt	cca	tcc	384
Cys	Gln	Ser	Trp	Thr	Cys	Pro	Tyr	Thr	Gly	Pro	Val	Ser	Ser	Pro	Ser	
			115				120					125				
tgg	aag	ttt	cac	tca	gat	gta	aat	ttc	acc	caa	gaa	gtc	agc	caa	gtg	432
Trp	Lys	Phe	His	Ser	Asp	Val	Asn	Phe	Thr	Gln	Glu	Val	Ser	Gln	Val	
	130					135					140					
tcc	ctt	cga	cta	cac	ttc	tct	aag	tgc	ggc	tcc	tcc	atg	acc	ctt	cta	480
Ser	Leu	Arg	Leu	His	Phe	Ser	Lys	Cys	Gly	Ser	Ser	Met	Thr	Leu	Leu	
145					150					155					160	
gta	gat	gcc	cct	gga	tat	gat	cct	tta	tgg	ttc	atc	acc	tca	gaa	ccc	528
Val	Asp	Ala	Pro	Gly	Tyr	Asp	Pro	Leu	Trp	Phe	Ile	Thr	Ser	Glu	Pro	
				165					170					175		
act	cag	cct	ccc	cca	act	cct	ccc	cca	ctg	gtc	cat	gac	tcc	gac	ctt	576
Thr	Gln	Pro	Pro	Pro	Thr	Pro	Pro	Pro	Leu	Val	His	Asp	Ser	Asp	Leu	
			180					185					190			
gaa	cac	gtc	cta	acc	ccc	tcc	acg	tct	tgg	aca	acc	aaa	atg	ctc	aag	624
Glu	His	Val	Leu	Thr	Pro	Ser	Thr	Ser	Trp	Thr	Thr	Lys	Met	Leu	Lys	
		195					200					205				
ttt	atc	cag	ctg	acc	ttg	cag	agc	acc	aat	tac	tcc	tgc	atg	gtt	tgc	672
Phe	Ile	Gln	Leu	Thr	Leu	Gln	Ser	Thr	Asn	Tyr	Ser	Cys	Met	Val	Cys	
	210					215					220					
gtg	gat	aga	tcc	agc	ctc	tca	tcc	tgg	cat	gtg	ctc	tac	acc	ccc	aac	720
Val	Asp	Arg	Ser	Ser	Leu	Ser	Ser	Trp	His	Val	Leu	Tyr	Thr	Pro	Asn	
225					230					235					240	
atc	tcc	att	ccc	caa	caa	acc	tcc	tcc	cga	acc	atc	ctc	ttt	cct	tct	768
Ile	Ser	Ile	Pro	Gln	Gln	Thr	Ser	Ser	Arg	Thr	Ile	Leu	Phe	Pro	Ser	
				245					250					255		
ctt	gcc	ctg	ccc	gct	cct	cca	ttc	caa	ccc	ttc	cct	tgg	acc	cat	tgc	816
Leu	Ala	Leu	Pro	Ala	Pro	Pro	Phe	Gln	Pro	Phe	Pro	Trp	Thr	His	Cys	
			260					265					270			
tac	caa	cct	cgc	cta	cag	gca	ata	acg	aca	gat	gac	tgc	aac	aac	tcc	864
Tyr	Gln	Pro	Arg	Leu	Gln	Ala	Ile	Thr	Thr	Asp	Asp	Cys	Asn	Asn	Ser	
		275				280						285				
att	atc	ctc	ccc	cct	ttt	tcc	ctc	gcc	ccc	gta	cct	cct	ccg	gcg	aca	912

Ile Ile Leu Pro Pro Phe Ser Leu Ala Pro Val Pro Pro Pro Ala Thr  
 290 295 300

<210> 45  
 <211> 304  
 <212> PRT  
 <213> Human T-cell lymphotropic virus type 2

<400> 45

Met Gly Asn Val Phe Phe Leu Leu Leu Phe Ser Leu Thr His Phe Pro  
 1 5 10 15

Pro Val Gln Gln Ser Arg Cys Thr Leu Thr Val Gly Ile Ser Ser Tyr  
 20 25 30

His Ser Ser Pro Cys Ser Pro Thr Gln Pro Val Cys Thr Trp Asn Leu  
 35 40 45

Asp Leu Asn Ser Leu Thr Thr Asp Gln Arg Leu His Pro Pro Cys Pro  
 50 55 60

Asn Leu Ile Thr Tyr Ser Gly Phe His Lys Thr Tyr Ser Leu Tyr Leu  
 65 70 75 80

Phe Pro His Trp Ile Lys Lys Pro Asn Arg Gln Gly Leu Gly Tyr Tyr  
 85 90 95

Ser Pro Ser Tyr Asn Asp Pro Cys Ser Leu Gln Cys Pro Tyr Leu Gly  
 100 105 110

Cys Gln Ser Trp Thr Cys Pro Tyr Thr Gly Pro Val Ser Ser Pro Ser  
 115 120 125

Trp Lys Phe His Ser Asp Val Asn Phe Thr Gln Glu Val Ser Gln Val  
 130 135 140

Ser Leu Arg Leu His Phe Ser Lys Cys Gly Ser Ser Met Thr Leu Leu  
 145 150 155 160

Val Asp Ala Pro Gly Tyr Asp Pro Leu Trp Phe Ile Thr Ser Glu Pro  
 165 170 175

Thr Gln Pro Pro Pro Thr Pro Pro Pro Leu Val His Asp Ser Asp Leu  
 180 185 190

Glu His Val Leu Thr Pro Ser Thr Ser Trp Thr Thr Lys Met Leu Lys  
 195 200 205

Phe Ile Gln Leu Thr Leu Gln Ser Thr Asn Tyr Ser Cys Met Val Cys  
 210 215 220

Val Asp Arg Ser Ser Leu Ser Ser Trp His Val Leu Tyr Thr Pro Asn  
 225 230 235 240

Ile Ser Ile Pro Gln Gln Thr Ser Ser Arg Thr Ile Leu Phe Pro Ser  
 245 250 255

Leu Ala Leu Pro Ala Pro Pro Phe Gln Pro Phe Pro Trp Thr His Cys  
 260 265 270

Tyr Gln Pro Arg Leu Gln Ala Ile Thr Thr Asp Asp Cys Asn Asn Ser  
 275 280 285

Ile Ile Leu Pro Pro Phe Ser Leu Ala Pro Val Pro Pro Pro Ala Thr  
 290 295 300

<210> 46

<211> 930

<212> ADN

<213> Simian T-cell lymphotropic virus type 3

<220>

<221> CDS

<222> (1) .. (930)

<223>

<400> 46

atg ggt aag ttt ggc ctt tat tgt ctt gtt cac ctt tac ata ctt ctc 48  
 Met Gly Lys Phe Gly Leu Tyr Cys Leu Val His Leu Tyr Ile Leu Leu  
 1 5 10 15

cct gcc tcc tct ggc aat ccc agt cgg tgc acc ctg ttc ata ggg gcc 96  
 Pro Ala Ser Ser Gly Asn Pro Ser Arg Cys Thr Leu Phe Ile Gly Ala  
 20 25 30

tct tcc tac cac tcc agc cct tgc ggg tcc agc ctc cca cgg tgt acc 144  
 Ser Ser Tyr His Ser Ser Pro Cys Gly Ser Ser Leu Pro Arg Cys Thr  
 35 40 45

tgg aat ctt gac cta ttc tcc ctc acg aaa gat caa agc cta agc ccc 192  
 Trp Asn Leu Asp Leu Phe Ser Leu Thr Lys Asp Gln Ser Leu Ser Pro  
 50 55 60

cca tgt cca gac tta att act tac tca caa tac cac aag ccc tac tcc 240  
 Pro Cys Pro Asp Leu Ile Thr Tyr Ser Gln Tyr His Lys Pro Tyr Ser  
 65 70 75 80

ctg tat gta ttc cct cat tgg ata act aaa cct aac cgc cgg ggc tta 288  
 Leu Tyr Val Phe Pro His Trp Ile Thr Lys Pro Asn Arg Arg Gly Leu  
 85 90 95

30/32

ggt tac tat tcc gct tcc tac tca gac ccc tgt gcc ata cag tgc cct	336
Gly Tyr Tyr Ser Ala Ser Tyr Ser Asp Pro Cys Ala Ile Gln Cys Pro	
100 105 110	
tac ctg gga tgc cag tgc tgg aca tgc ccc tat acg ggc ccg gtg tcc	384
Tyr Leu Gly Cys Gln Ser Trp Thr Cys Pro Tyr Thr Gly Pro Val Ser	
115 120 125	
agt ccg cat tgg aga tac acc tat gat ctt aac ttt acc cag gag gta	432
Ser Pro His Trp Arg Tyr Thr Tyr Asp Leu Asn Phe Thr Gln Glu Val	
130 135 140	
tca tcc gtc tcc tta cac ttg cat ttc tcc aaa tgc gga tcc tgc ttc	480
Ser Ser Val Ser Leu His Leu His Phe Ser Lys Cys Gly Ser Ser Phe	
145 150 155 160	
tcc ttt cta cta gac gca cca gga tat gac cca gtg tgg ttc ctc tcc	528
Ser Phe Leu Leu Asp Ala Pro Gly Tyr Asp Pro Val Trp Phe Leu Ser	
165 170 175	
tcc cag gcc aca cag gct cca ccc aca cct gcc cct ctc ata cgg gac	576
Ser Gln Ala Thr Gln Ala Pro Pro Thr Pro Ala Pro Leu Ile Arg Asp	
180 185 190	
tca gat ctc cag tac att cta gaa ccg ccc att ccg tgg agc tct aag	624
Ser Asp Leu Gln Tyr Ile Leu Glu Pro Pro Ile Pro Trp Ser Ser Lys	
195 200 205	
att ctt aac ctt atc ctc ctc acc cta aaa agc act aac tat tct tgc	672
Ile Leu Asn Leu Ile Leu Leu Thr Leu Lys Ser Thr Asn Tyr Ser Cys	
210 215 220	
atg gtc tgt gtt gac cgc tcc agc cta tcc tca tgg cat gtc ctg tat	720
Met Val Cys Val Asp Arg Ser Ser Leu Ser Ser Trp His Val Leu Tyr	
225 230 235 240	
gga ccc act caa gtc ccc agt cca ccc gac ccc caa gcc cgg tct atc	768
Gly Pro Thr Gln Val Pro Ser Pro Pro Asp Pro Gln Ala Arg Ser Ile	
245 250 255	
ctg cga cct gcc tta gct att ccc gcc agt aat atc acc ccc ccg ttt	816
Leu Arg Pro Ala Leu Ala Ile Pro Ala Ser Asn Ile Thr Pro Pro Phe	
260 265 270	
cct tgg acc cat tgc tat cgc cct cct ccg caa gcc atc tcc tgc gag	864
Pro Trp Thr His Cys Tyr Arg Pro Pro Pro Gln Ala Ile Ser Ser Glu	
275 280 285	
aat tgt aac aac tct gta gtg ctg ccc ccc ttt tct ctg tct cca att	912
Asn Cys Asn Asn Ser Val Val Leu Pro Pro Phe Ser Leu Ser Pro Ile	
290 295 300	
cct aac gtc tcc aga ccc	930
Pro Asn Val Ser Arg Pro	
305 310	

&lt;210&gt; 47

&lt;211&gt; 310

&lt;212&gt; PRT

&lt;213&gt; Simian T-cell lymphotropic virus type 3

&lt;400&gt; 47

Met Gly Lys Phe Gly Leu Tyr Cys Leu Val His Leu Tyr Ile Leu Leu  
 1 5 10 15

Pro Ala Ser Ser Gly Asn Pro Ser Arg Cys Thr Leu Phe Ile Gly Ala  
 20 25 30

Ser Ser Tyr His Ser Ser Pro Cys Gly Ser Ser Leu Pro Arg Cys Thr  
 35 40 45

Trp Asn Leu Asp Leu Phe Ser Leu Thr Lys Asp Gln Ser Leu Ser Pro  
 50 55 60

Pro Cys Pro Asp Leu Ile Thr Tyr Ser Gln Tyr His Lys Pro Tyr Ser  
 65 70 75 80

Leu Tyr Val Phe Pro His Trp Ile Thr Lys Pro Asn Arg Arg Gly Leu  
 85 90 95

Gly Tyr Tyr Ser Ala Ser Tyr Ser Asp Pro Cys Ala Ile Gln Cys Pro  
 100 105 110

Tyr Leu Gly Cys Gln Ser Trp Thr Cys Pro Tyr Thr Gly Pro Val Ser  
 115 120 125

Ser Pro His Trp Arg Tyr Thr Tyr Asp Leu Asn Phe Thr Gln Glu Val  
 130 135 140

Ser Ser Val Ser Leu His Leu His Phe Ser Lys Cys Gly Ser Ser Phe  
 145 150 155 160

Ser Phe Leu Leu Asp Ala Pro Gly Tyr Asp Pro Val Trp Phe Leu Ser  
 165 170 175

Ser Gln Ala Thr Gln Ala Pro Pro Thr Pro Ala Pro Leu Ile Arg Asp  
 180 185 190

Ser Asp Leu Gln Tyr Ile Leu Glu Pro Pro Ile Pro Trp Ser Ser Lys  
 195 200 205

Ile Leu Asn Leu Ile Leu Leu Thr Leu Lys Ser Thr Asn Tyr Ser Cys  
 210 215 220

Met Val Cys Val Asp Arg Ser Ser Leu Ser Ser Trp His Val Leu Tyr  
 225 230 235 240

Gly Pro Thr Gln Val Pro Ser Pro Pro Asp Pro Gln Ala Arg Ser Ile  
245 250 255

Leu Arg Pro Ala Leu Ala Ile Pro Ala Ser Asn Ile Thr Pro Pro Phe  
260 265 270

Pro Trp Thr His Cys Tyr Arg Pro Pro Pro Gln Ala Ile Ser Ser Glu  
275 280 285

Asn Cys Asn Asn Ser Val Val Leu Pro Pro Phe Ser Leu Ser Pro Ile  
290 295 300

Pro Asn Val Ser Arg Pro  
305 310